

الإحصاء التحليلي

مع تطبيقات برمجية

SPSS

الدكتور
نبيل جمعة صالح النجار





mohamed khatab

الإحصاء التحليلي
مع تطبيقات برمجية SPSS

الإحصاء التحليلي

مع تطبيقات برمجية SPSS

د. نبيل جمعه صالح النجار

دكتوراه الفلسفة في التربية

قسم القياس والتقويم



محفوظ جميع الحقوق

رقم التصنيف : 519.50.285
المؤلف ومن هو في حكمه : النجار، نبيل جمعة صالح.
عنوان الكتاب : الاحصاء التحليلي مع تطبيقات برمجية SPSS.
رقم الإيداع : 2014/6/2724
الواصفات : /الاحصاء//الحواسيب/
بيانات الناشر : عمان - دار ومكتبة الحامد للنشر والتوزيع
يحمل المؤلف كامل المسؤولية القانونية عن محتوى مصنفه ولا يعبر هذا المصنف عن رأي دائرة المكتبة الوطنية أو أي جهة حكومية أخرى.
(ردمك) ISBN 978-9957-32-873-3

تم إعداد بيانات الفهرسة والتصنيف الأولية من قبل دائرة المكتبة الوطنية.

لا يجوز نشر أو اقتباس أي جزء من هذا الكتاب، أو اختزان مادته بطريقة الاسترجاع، أو نقله على أي وجه، أو بأي طريقة
أكانت إلكترونية، أم ميكانيكية، أم بالتصوير، أم التسجيل، أم بخلاف ذلك، دون الحصول على إذن المؤلف الخطي، وبخلاف ذلك
يتعرض الفاعل للملاحقة القانونية.

الطبعة الأولى 2015-1436هـ



دار الحامد للنشر والتوزيع

الأردن - عمان - شفا بدران - شارع العرب مقابل جامعة العلوم التطبيقية

هاتفه +962 6 5231081 فاكس : +962 6 5235594

ص.ب. (366) الرمز البريدي (11941) عمان - الأردن

www.daralhamed.net

E-mail : daralhamed@yahoo.com

المحتويات

الصفحة	الموضوع
11	المقدمة
	الفصل الأول
13	مفاهيم أساسية في الإحصاء Statistical Concepts
15	1-1 المفاهيم الإحصائية Statistical Concepts
20	2-1 مقاييس النزعة المركزية Measures of Central Tendency الوسط الحسابي، الوسيط، المنوال، المئينات
	3-1 مقاييس التشتت Measure Dispersion or Variation
26	المدى، الانحراف المتوسط، التباين، الانحراف المعياري، معامل الاختلاف
	4-1 الارتباط Correlation والانحدار Regression
33	الانحدار ومفهومه، الانحدار الخطي البسيط، معادلة خط الانحدار
39	5-1 استخدام برمجية SPSS في حل المسائل.
61	6-1 تمارين Exercise.
	الفصل الثاني
73	التوزيعات المجتمعية الاحتمالية والتوزيعات العينية Probability Distributions & Sampling Distributions
75	1-2 المقدمة Introduction
76	2-2 التوزيع الطبيعي Normal Distribution
83	3-2 التوزيع الاحتمالي التائي Probability Distribution (T)
85	4-2 توزيع مربع كاي χ^2 Probability Distribution
86	5-2 التوزيع الفائي F Probability Distribution

88	Sampling Theory	6-2 نظرية المعاينة
90	Samples	7-2 العينات
101	Sampling Distribution	8-2 توزيع المعاينة
109	Exercise	9-2 أسئلة وتمارين

الفصل الثالث

التقدير وفترات الثقة

113	تقدير معالم المجتمع من معالم العينة (باستخدام عينة واحدة) Estimation of Population Parameters
-----	--

115	1-3 مقدمة Introduction
116	2-3 خواص جودة التقدير Properties of Goodness of Estimation
116	3-3 أنواع القيم التقديرية لمعالم المجتمع.
117	4-3 فترات الثقة Confidence Interval
127	5-3 تمارين Exercise

الفصل الرابع

اختبار الفرضيات

Hypothesis Testing

131	اختبار الفرضيات
133	1-4 اختبار الفرضيات Hypothesis Testing
136	2-4 اختبار الفرضيات الإحصائية Testing Statistical Hypothesis
136	3-4 مفاهيم أساسية في فحص الفرضيات Basic Concepts in Hypothesis Testing
139	4-4 خطوات اختبار الفرضيات Hypothesis Testing Steps
140	5-4 استخدام برمجية SPSS في حل المسائل.
142	6-4 تمارين Exercise

الفصل الخامس

145 اختبار الفرضيات التي تتعلق بالمتوسطات الحسابية Testing Hypothesis Regarding Mean

- 147 1-5 اختبار الفرضية المتعلقة بوسط حسابي واحد (مجتمع واحد)
- 148 * اختبار الفرضية المتعلقة بوسط حسابي واحد (حجم العينة كبير، تباين المجتمع معلوم)
- 151 * (حجم العينة كبير، تباين المجتمع غير معلوم)
- 153 * (σ غير معلومة والعينة صغيرة الحجم)
- 155 2-5 اختبار الفرضيات حول الفرق بين وسطين حسابيين
- 155 * اختبار الفرضيات حول الفرق بين وسطين للبيانات المستقلة
- 158 * اختبار الفرضيات حول الفرق بين وسطين للبيانات غير المستقلة
- 164 3-5 استخدام برمجية SPSS في حل المسائل.
- 173 4-5 تمارين Exercise.

الفصل السادس

185 اختبار الفرضيات حول التباينات Testing Hypothesis Inference Regarding Variances

- 187 1-6 اختبار فرضية تتعلق بالتباين لمجتمع واحد.
- 191 2-6 اختبار فرضية تتعلق بتساوي التباين لمجتمعين مستقلين.
- 193 3-6 اختبار فرضية تتعلق بتساوي التباين لمجتمعين غير مستقلين.
- 195 4-6 تمارين Exercise.

الفصل السابع

199 اختبار الفرضيات حول معاملات الارتباط Hypotheses Testing Regarding Correlation Coefficients

- 201 1-7 مقدمة
- 205 2-7 اختبار الفرضية حول معامل ارتباط واحد.
- 206 3-7 اختبار الفرضيات حول الفرق بين معاملي ارتباط مستقلين.
- 207 4-7 اختبار الفرضيات حول معاملي ارتباط للبيانات غير المستقلة.

- 211 5-7 استخدام برمجية SPSS في حل المسائل.
- 221 6-7 تمارين Exercise.

الفصل الثامن

اختبار الفرضيات حول النسب

Hypothesis Testing Regarding Proportions

- 225 1-8 اختبار الفرضيات حول النسب.
- 227 2-8 اختبار الفرضيات حول نسبة واحدة .
- 230 3-8 اختبار الفرضيات حول نسبتي مستقلتين.
- 232 4-8 اختبار الفرضيات حول نسبتي للبيانات الغير مستقلة
- 234 5-8 اختبار χ^2 لحسن المطابقة The Chi-Square Goodness of fit
- 235 6-8 استخدام برمجية SPSS في حل المسائل.
- 237 7-8 تمارين Exercise.

الفصل التاسع

تحليل التباين

Analysis of Variance

- 239 1-9 مقدمة.
- 243 2-9 تحليل التباين الأحادي One-Way Analysis of Variance
- 256 3-9 تحليل التباين الثنائي Tow-Way Analysis of Variance
- 268 4-9 تحليل التباين (ANCOVA) Analysis of Covariance
- 281 5-9 استخدام برمجية SPSS في حل المسائل.
- 288 6-9 تمارين Exercise.

الفصل العاشر

المقارنات المتعددة

Multiple Comparisons

- 303 1-10 مقدمة.
- 303 2-10 أنواع المقارنات المتعددة
- 303 1- المقارنات المخطط لها

304	أ- طريقة المقارنات المتعامدة Orthogonal
305	ب- طريقة دن Dunn وتسمى أيضاً طريقة بنفوروني Bonferroni
306	2- المقارنات غير المخطط لها
306	* طريقة شافيه Scheffe، طريقة توكي Tukey،
307	* طريقة توكي Tukey
307	* طريقة نيومان كولز Newman Kuelz
310	3-10 تمارين Exercise.

الفصل الحادي عشر التحليل العاملي Factor Analysis

311	
313	1-11 مقدمة
317	2-11 مفهوم التحليل العاملي
319	3-11 أهمية التحليل العاملي ومبادئه
320	4-11 أهداف التحليل العاملي
322	5-11 طرق التحليل العاملي
332	6-11 بعض مفاهيم التحليل العاملي
335	7-11 استخدام برمجية SPSS في حل المسائل
347	8-11 تمارين Exercise

الفصل الثاني عشر الاحصاءات الالامعلمية NONPARAMETRIC STATISTICS

349	
351	1-12 مقدمة Introduction
352	2-12 الطرق الالامعلمية (عينة واحدة) (Single Sample)
354	3-12 الطرق الالامعلمية (عينتين مستقلتين) Nonparametric Methods (Two Independent Samples)
358	4-12 الطرق الالامعلمية (عينتين مرتبطتين) Nonparametric Methods (Two Related Samples)

- 361 5-12 الطرق الالاعلمية (ثلاثة عينات مستقلة أو أكثر)
Nonparametric Methods(3 or more Independent Samples)
- 365 6-12 الطرق الالاعلمية (ثلاثة عينات مرتبطة أو أكثر)
Nonparametric Methods (3 or more Related Samples)
- 367 7-12 استخدام برمجية SPSS في حل المسائل.
- 374 8-12 تمارين Exercise.

الفصل الثالث عشر برنامج التحليل الإحصائي SPSS

- 381 1-13 التعرف على بيئة النظام الإحصائي SPSS
- 382 2-13 تشغيل نظام SPSS.
- 382 3-13 شاشات نظام SPSS.
- 384 4-13 ملفات نظام SPSS.
- 385 5-13 القوائم الرئيسية لبرنامج SPSS.
- 391 6-13 استخدام برمجية SPSS في حل المسائل.

الملاحق

- 413 الملحق 1: جداول التوزيعات الاحتمالية
- 415 جدول Z
- 416 جدول T
- 417 جدول F
- 418 جدول χ^2

المصادر

- 419 - المصادر العربية
- 421 - المصادر الأجنبية

مقدمة

الحمد لله رب العالمين والصلاة والسلام على سيدنا محمد صلى الله عليه وسلم
خير الأنبياء المرسلين.

لقد أطلق على هذا العصر عصر المعلوماتية، إن هذا العصر هو عصر تكنولوجيا
المعلومات، فقد أصبحت المعلومات تشكل الثورة الحقيقية التي سادت في الحياة العصرية ودخلت
إلى بيوتنا وأعمالنا بشكل أو بآخر شئنا ذلك أم أبينا، لذا يجب علينا أن نساير ونواكب التقدم
في شتى المجالات والميادين.

وبما أن الإحصاء علم مهم للكثير من الطلاب والباحثين فقد تم تغطية جميع فصول هذا
الكتاب بتطبيق ما ورد فيها من معلومات نظرية من خلال التمارين العملية وعن طريق استخدام
برمجية SPSS.

لقد راعى المؤلف أن يكون هذا الكتاب متوافقاً مع خطة منهاج الإحصاء التحليلي
للبحوث التربوية" المقرر في كلية العلوم التربوية في جامعة مؤتة، ومع المستوى المطلوب لمادة
"الإحصاء التحليلي" و"الإحصاء المتقدم" على مستوى الجامعات الحكومية والخاصة عموماً،
مع الأخذ بعين الاعتبار أن يكون مناسباً لجميع من لديه حب المعرفة بعلم الإحصاء التحليلي
وتطبيقاته وفوائده وبطريقة سهلة وبسيطة.

ويأتي هذا الكتاب كأحد الوسائل التي توفر لطلبة الدراسات العليا على مستوى
الجامعات فرصة تمكنه من تعلم أساسيات مادة الإحصاء التحليلي". وذلك من خلال ما يحويه
من مواضيع جديدة وهامة في مفاهيم ومبادئ الإحصاء التحليلي وتطبيقاته باستخدام برمجية
SPSS.

د. نيل جمع النجار

Email: nabilnajar@yahoo.com

Mobile: 0777757837

Mobile: 0785623442

Mobile: 0798011404

الفصل الأول

مفاهيم أساسية في الإحصاء

Statistical Concepts

- 1-1 المفاهيم الإحصائية Statistical Concepts
مفهوم الإحصاء، الرموز الإحصائية، المتغيرات.
- 2-1 مقاييس النزعة المركزية Measures of Central Tendency
الوسط الحسابي، الوسيط، المنوال، المئينات.
- 3-1 مقاييس التشتت Measure Dispersion or Variation
المدى، الانحراف المتوسط، التباين، الانحراف المعياري، معامل الاختلاف.
- 4-1 الارتباط Correlation والانحدار Regression
أنواع الارتباط، قياس الارتباط.
الانحدار ومفهومة، الانحدار الخطي البسيط، معادلة خط الانحدار.
- 5-1 استخدام برمجية SPSS في حل المسائل.
- 6-1 تمارين Exercise.

الفصل الأول

مفاهيم أساسية في الإحصاء Statistical Concepts

1-1 المفاهيم الإحصائية Statistical Concepts

Definition of Statistics مفهوم الإحصاء

بمجموعة الطرق العلمية التي تعنى بجمع وتصنيف وتبويب وتفسير وتلخيص وتقييم البيانات والخروج منها باستنتاجات حول المجتمع من خلال اعتماد جزء صغير من المجتمع (العينة).

الإحصاء الوصفي Descriptive Statistics

جمع المعطيات وتحليلها ووصفها وإظهارها بصيغة مفهومة وذات مدلول والتعامل مع المعطيات الإحصائية من دون تعميم، وعرضها عن طريق الجداول والرسوم البيانية وغيرها.

الإحصاء الاستدلالي Inferential Statistics

تحليل وتفسير وتقدير واستخلاص الاستنتاجات بالاعتماد على عينة من المجتمع للتوصل إلى قرارات تخص المجتمع ويتعامل مع التعميم والتنبؤ.

المعطيات الإحصائية Statistical Data

البيانات والمعلومات الإحصائية المتعلقة بالظواهر الإدارية والاجتماعية والترهوية وتختلف المعطيات من حيث نوعها وطبيعتها باختلاف الظاهرة المطلوب قياسها وباختلاف منهجية البحث والأدوات الإحصائية المستخدمة.

البيانات Data: مشاهدات، علامات، مقادير، يعبر عنها بأرقام.

البيانات غير المبوبة Ungrouped Data: البيانات الأولية أو الأصلية التي جمعت ولم

تبويب.

البيانات المبوبة **Grouped Data**: البيانات التي بوبت وفرغت في جدول توزيع تكراري.

المعطيات الكمية **Quantitative Data**

تصف الظاهرة بشكل رقمي عن ظاهرة معينة، مثل علامة الطالب وسعر السلعة.

المعطيات النوعية **Qualitative Data**

تصف الظاهرة المعنية بشكل غير رقمي عن ظاهرة معينة، مثل الجنس، اللون.

التوزيع **Distribution**: مجموعة مشاهدات مهما كان عددها.

المجتمع **Population**: مجتمع بيانات أو مشاهدات أو علامات يحدد هويته الباحث.

مجتمع العينة **Sample Population**: المجتمع الذي تؤخذ منه العينة.

مجتمع الهدف **Target Population**: المجتمع الذي ستعمم عليه نتائج الدراسة التي أجريت على مجتمع العينة.

العينة **Sample**: مجموعة جزئية من المجتمع.

المؤشر **Index**: تدل على جميع مقاييس الرعة المركزية والتشتت والعلاقة (الارتباط) سواء محسوبة لعينات أو لمجتمعات.

مؤشر عينة (إحصائي) **Statistic**: يستخدم للعينات، مثل الوسط الحسابي لعينة X' .

مؤشر مجتمع (معلم) **Parameter**: يستخدم للمجتمع، مثل الوسط الحسابي للمجتمع μ .

لكل توزيع خصائص (توصيف) هي: مقاييس الرعة المركزية.

* الرموز الإحصائية **Statistical Symbols**

في الإحصاء الاستدلالي تستخدم إحصاءات العينة كتقديرات لمعالم المجتمع المناظر.

رموز العينة	المعنى	رموز المجتمع	المعنى
X'	متوسط عينة	μ_x	متوسط مجتمع.
S^2_x	تباين عينة	σ^2_x	تباين مجتمع.

انحراف معياري لعينة	S_x	انحراف معياري لمجتمع	σ_x
معامل ارتباط لعينة	r_{xy}	معامل ارتباط لمجتمع	ρ_{xy}
نسبة عينة	P	نسبة المجتمع	π
عدد مشاهدات العينة	n	عدد مشاهدات المجتمع	N
عدد المشاهدات	n	Number of Scores	
مجموع المشاهدات	Σx	Sum of x	
الوسيط	M_d	Median	
الوسط الحسابي للعينة	\bar{X}	Sample Mean of X's	
الوسط الحسابي للمجتمع	μ	Mu; Population Mean	
علامة خام	X, Y	Raw Score	
التكرار	f	Frequency	
التكرار التراكمي	cf	Cumulative Frequency	
التكرار النسبي	$rel.f$	Relative Frequency	
درجات الحرية	df	Degree Freedom	
انحراف معياري لعينة	S_x	Sample Standard Deviation	
تباين عينة	S^2_x	Sample Variance	
انحراف معياري لمجتمع	σ_x	Population Standard Deviation	
تباين مجتمع	σ^2_x	Population Variance	
معامل ارتباط لمجتمع	ρ	Population Correlation Coefficient	
معامل ارتباط لعينة	r	Sample Correlation Coefficient	
العلامة المعيارية	Z	Standard Score	
الخطأ المعياري للوسط	$\sigma_{x'}$	Standard Error of Mean	
الفرق بين البيانات	D	Difference Score	$(X-Y)$
معامل ارتباط بيرسون	r	Pearson Correlation Coefficient	
معامل ارتباط سبيرمان	r_s	Sperman Correlation Coefficient	

Point-biserial Correlation Coefficient	ر بويت بايسير يال	r_{pb}
Regression Line Equation	معادلة خط الانحدار	$Y=ax+b$
Slope of the Regression Line	ميل خط الانحدار	b
Y-intercept of the Regression Line	نقطة تقاطع خط الانحدار	a
Coefficient of Determination	معامل التحديد	\bar{D}
Mean of Difference Scores ($\sum D/n$)	الوسط الحسابي للفروق	

* المتغيرات Variables

المتغيرات Variables ما إحصائية أو عشوائية فالمتغير الإحصائي يمثل القيم التي تأخذها ظاهرة ما، والمتغير العشوائي عبارة عن ظاهرة نوعية أو كمية لا يمكن التنبؤ بها بشكل مسبق. المتغير Variable: ظاهرة تظهر اختلافات بين قيمها، إذا اختلفت الخاصية عند أفراد مجموعة معينة كماً أو نوعاً نقول بأنها هي المتغير، إذا كان الأفراد متساوين كماً أو متشابهين نوعاً بالنسبة لخاصية معينة تكون هي الثابت.

* تصنف المتغيرات حسب طبيعة المعلومات التي يؤديها القياس إلى:

1- المتغيرات الاسمية Nominal Variables

المتغيرات النوعية التي لها عدد فئات محدد من دون أي وزن لهذه الفئات ولا يوجد أفضلية لأحدها على الآخر، وتستخدم لغايات التصنيف فقط. مثال: متغير الجنس ويصنف فيه المجتمع إلى فئتين هما الذكور والإناث فلو رمزنا للذكور بالرقم (1) والإناث بالرقم (2) فالأرقام ليس لها معنى حقيقي ولا يمكن إجراء العمليات الحسابية عليها.

مثال: إذا قسم الأفراد حسب الطول إلى طويل وقصير.

مثال: أوجه قطعة النقد وهي صورة وكتابة.

2- المتغيرات الترتيبية Ordinal Variables

متغير نوعي ذو عدد محدد من الفئات يمكن ترتيبها تصاعدياً أو تنازلياً، ولا يمكن تحديد الفروق بدقة بين القيم المختلفة. مثال: كبير وسط صغير.

مثال: A أكبر من B ولكن لا نستطيع معرفة كم يكبر A عن B.

مثال: إذا كانت علامة جمال في مادة اللغة العربية أكثر من علامة فادي، وأن علامة فادي أكثر من علامة نبيل، فإننا نعرف هنا ترتيب الأفراد فقط.

3- المتغيرات الفئوية Interval Variables

المتغيرات الكمية التي يمكن إجراء العمليات الحسابية عليها وذلك دون أن تتأثر المسافة النسبية بين قيمها، ويميز هذا المتغير من خلال قيمة الصفر التي لا تعني عدم توافر تلك الصفة.

مثال: إذا كانت علامة جمال في مادة اللغة العربية (60) وعلامة أيمن (50)، وعلامة محمد (40)، فإننا نعرف الترتيب، ونعرف كم تزيد علامة كل طالب عن الآخر.

مثال: إذا حصل محمد على علامة (صفر) في اختبار اللغة الإنجليزية فهذا لا يعني أن محمد لا يعرف شيئاً في اللغة الإنجليزية.

مثال: إذا كانت درجة الحرارة تساوي (صفر) فهذا لا يعني عدم وجود درجة حرارة.

4- المتغيرات النسبية Ratio Variables

متغيرات كمية ليس لها فئات محددة وهي تشبه المتغيرات الفئوية ولكن الصفر هنا يمثل عدم توفر الصفة، ومثال له المتغيرات الزمنية.

مثال: علامات الطلاب في مساق معين، أوزان الطلاب.

المقياس	الاستخدام
1- الاسمي Nominal	تصنيف
2- الترتيبي Ordinal	تصنيف + ترتيب
3- الفئوي Interval	تصنيف + ترتيب + مسافة + صفر افتراضي
4- النسبي Ratio	تصنيف + ترتيب + مسافة + صفر مطلق

* تصنف المتغيرات حسب وجود علاقة بين متغيرين إلى:

1. المتغير المستقل Independent Variable: وهو المتغير الذي يخضع لسيطرة الإحصائي أو الباحث.

2. المتغير التابع Dependent Variable: وهو المتغير الذي نتبأ بقيمته من خلال معرفتنا لقيم المتغير المستقل.

مثال: إذا أراد مدرس أن يبحث عن أثر عدد ساعات الدراسة على تحصيل الطالب في مبحث معين، حدد المتغير المستقل والمتغير التابع؟

المتغير المستقل: عدد ساعات الدراسة. المتغير التابع: تحصيل الطالب.

مثال: أثر الغياب على تحصيل الطالب.

المتغير المستقل: الغياب. المتغير التابع: تحصيل الطالب.

2-1 مقاييس النزعة المركزية

Measures of Central Tendency

مقدمة:

لأي بيانات إحصائية هناك خواص تساعد على إعطاء فكرة ومدلول عن وضع هذه البيانات ومن هذه الخصائص ما يلي:

- إحصائيات النزعة المركزية وتسمى بالمتوسطات وأهم هذه المتوسطات هي: الوسط الحسابي والوسيط والمنوال، ونتمكن بواسطتها من تحديد موقع النقطة التي تتمحور حولها كافة القيم، ولكل من هذه الإحصائيات مزايا وعيوب، ويمكن استخدام أي منها بناء على عدة أمور منها:

• شكل التوزيع: هل هو معتدل أم ملتو.

• مستوى القياس: هل هو اسمي، رتي، فتوي، نسبي.

- إحصائيات التشتت: ويقصد بها حالة الانتشار التي تكون عليها البيانات حول المتوسط، وأهم هذه المقاييس المدى والتباين والانحراف المعياري.

1- الوسط الحسابي (\bar{X}) Arithmetic Mean

الوسط الحسابي هو معدل المشاهدات في التوزيع.
الوسط الحسابي هو مجموع قيم المشاهدات على عددها.

* الوسط الحسابي للبيانات غير المبوبة Ungrouped Data

$$\bar{X} = \frac{\sum X_i}{n} \quad \text{الوسط الحسابي} \quad \text{س} = \frac{\text{مجموع المشاهدات}}{\text{عددها}}$$

حيث أن:

X : رمز الوسط الحسابي للعينة ويقرأ أكس بار.

\sum : رمز المجموع ويقرأ سيجمما.

X_i : رمز المشاهددة.

n : عدد المشاهدات في العينة.

* الوسط الحسابي للبيانات المبوبة Grouped Data

$$\bar{X} = \frac{\sum X_i \cdot F_i}{\sum F_i}$$

$$\bar{X} = \frac{\sum X_i \cdot F_i}{\sum F_i}$$

X_i : مراكز الفئات.

F_i : التكرار.

* الدلالات الإحصائية للوسط الحسابي Statistical Significance for Mean

- كلما ارتفعت قيمة الوسط الحسابي للعلامات دل ذلك على أداء أفضل، بشرط أن لا تكون هناك قيم متطرفة عالية أدت إلى ارتفاع الوسط الحسابي .
- كلما كانت العلامات موزعة على جانبي وسطها الحسابي بشكل متماثل ومتساو كان التوزيع معتدلاً وكاشفاً عن الفروق بين الطلاب بصورة أفضل.

* خصائص الوسط الحسابي Mean's Characteristics

- يعتمد على جميع المشاهدات.
- سهل الفهم والتفسير، ويتم حسابه بسهولة وسرعة.
- يتأثر بالتحويلات الخطية، ولا يتأثر باختلاف العينات في المجتمع.
- مقياس النزعة المفضل عند الحديث عن الإحصاء الاستدلالي.
- مجموع انحرافات القيم عن وسطها الحسابي يساوي صفراً.

* عيوب الوسط الحسابي Mean's Deficiencies

- لا يمكن قياسه بالطرق البيانية.
- قابليته للتأثر بعدد قليل من المشاهدات المتطرفة.
- لا يمكن حسابه في التوزيع ذي الفئات المفتوحة.

2- الوسيط (Median (M_d)

المشاهدة التي تقسم التوزيع إلى نصفين بحيث يكون فوقها 50% من المشاهدات ودونها 50% من المشاهدات .

الوسيط: المشاهدة التي تقع في منتصف التوزيع. وهو المئين 50 (P50).

حساب الوسيط:

- نعمل على ترتيب البيانات تصاعدياً أو تنازلياً.
- إذا كان عدد البيانات فردياً تكون المشاهدة التي ترتيبها $(n+1)/2$ هي الوسيط.
- إذا كان عدد البيانات زوجياً يكون معدل المشاهدين اللتان ترتيبهما $(n/2)$ و $(n/2)+1$ هو الوسيط.

* الوسيط للبيانات المبوبة Grouped data

$$و = \frac{\text{الحد الأدنى الفعلي لفئة الوسيط} + (ت_r - ت_{r-1}) \times \frac{ت_r - ت_{r-1}}{ت_r - ت_{r-1}}}{ت_r - ت_{r-1}}$$

ت_ر : التكرار التراكمي للوسيط.

ت_ص : التكرار التراكمي الصاعد للفئة السابقة لفئة الوسيط.

ت_ر : التكرار العادي لفئة الوسيط.

ط : طول الفئة.

$$M_d = L + \frac{(\sum f_i/2) - f_1}{f_2 - f_1} * H$$

حيث أن:

L : الحد الأدنى الفعلي لفئة الوسيط. H : طول الفئة. $(\sum F_i / 2)$: قيمة موقع الوسيط.

f₁ : التكرار المتجمع السابق لموقع الوسيط f₂ : التكرار المتجمع اللاحق لموقع الوسيط

Median's Characteristics خصائص الوسيط

- يمكن احتسابه للجداول المفتوحة.
- يمكن احتسابه في حالة فقدان بعض القيم شرط ان يكون ترتيبها معروفاً.
- سهولة احتسابه، يمكن إيجادها بيانياً.
- لا يتأثر بالقيم المتطرفة، موقعه يتوسط البيانات.
- لا يعتمد على قيم البيانات وإنما يعتمد على موقعها.
- يتأثر بالتحويلات الخطية (الجمع والطرح والضرب والقسمة).

Median's Deficiencies عيوب الوسيط

- حساس للقيم الوسيطة.
- إذا كان عدد المشاهدات قليل فالوسيط قد لا يعبر بصورة واضحة صحيحة عن مركز تجمع المشاهدات.

3- المنوال (Mode) : Mode (M_o)

العلامة أو مركز الفئة أو الصفة التي تقابل أعلى تكرار في التوزيع.

القيمة الأكثر تكراراً أو الظاهرة الأكثر شيوعاً.

* المنوال للبيانات غير المبوبة Ungrouped Data

■ ليس من الضروري وجود منوال للبيانات.

* المنوال للبيانات المبوبة Grouped Data

1- طريقة الفروق (بيرسون):

– تحديد الفئة المتوالية: وهي الفئة التي يقابلها أكبر تكرار.

– استخدام الصيغة التالية لحساب قيمة المنوال

$$M_o = L + \frac{d_1}{d_1 + d_2} * H$$

H : طول الفئة.

حيث ان: L : الحد الأدنى الفعلي للفئة المتوالية.

d₁ : تكرار الفئة المتوالية – تكرار الفئة السابقة

d₂ : تكرار الفئة المتوالية – تكرار الفئة اللاحقة

2- طريقة الرافعة

$$M_o = L + \frac{d_a}{d_a + d_b} * H$$

حيث ان: L : الحد الأدنى الفعلي للفئة المتوالية. d_a : تكرار الفئة اللاحقة.

H : طول الفئة. d_b : تكرار الفئة السابقة.

$$\text{المنوال} = \frac{\text{الحد الأدنى الفعلي} + \frac{\text{تكرار الفئة اللاحقة}}{\text{تكرار الفئة اللاحقة} + \text{تكرار الفئة السابقة}} \times \text{طول الفئة}}{2}$$

خصائص المنوال Mode's Characteristics

■ محدود الاستعمال ، و يتأثر كثيراً بحجم العينة.

■ الإحصائي الوحيد الذي يمكن استعماله عندما تكون البيانات الإحصائية بمستوى

القياس الاسمي.

- يمكن استعماله عندما تكون البيانات رتبية أو فئوية أو نسبية.
 - لا يعتمد على جميع قيم البيانات وإنما يعتمد على القيم المتكررة.
 - يتأثر بطول الفئة في التوزيع، ويمكن إيجادها بيانياً.
 - يتأثر بالتحويلات الخطية (الجمع والطرح والضرب والقسمة).
 - لا يتأثر بالقيم المتطرفة.
 - قد يكون للتوزيع أكثر من منوال.
 - لا يدخل كثيراً في تحليلات إحصائية خارج نطاق وصف البيانات.
 - يمكن احتسابه للجداول المفتوحة.
 - أقل تعبيراً كمتوسط عندما تكون القيم منتشرة على مدى واسع.
 - يمكن أن يكون للبيانات أكثر من منوال واحد.
- مثال 4-11: البيانات التالية: 5، 10، 27، 32، 18، 10، 40، 32 ليس من الضروري وجود منوال للبيانات.
- المنوال يمكن أن يحسب للمتغيرات الكمية والمتغيرات النوعية.

4- المئينات (Percentiles (P_i :

مقياس يتم بموجبه تقسيم البيانات إلى 100 جزء متساوي وبالتالي يوجد 99 مئين. وهو قيمة معينة ضمن التوزيع تسبقها أو تليها نسبة مئوية معينة من المشاهدات الداخلة فيه. المئين ك: المشاهدة التي يقل عنها أو يساويها ك% من المشاهدات، ويرمز لمئين ك بالرمز P_K ، إن قيمة المئين i تحسب عن طريق الصيغة التالية:

$$P_i = L + \frac{[i \sum f_j / 100] - f_1}{f_2 - f_1} * H$$

* إيجاد المئين للبيانات غير المهوبة Ungrouped Data

تعريف: الرتبة المئينية لمشاهدة ما هي النسبة المئوية للتكرار التراكمي المقابل لتلك المشاهدة بالنسبة إلى مجموع التكرارات.

$$\text{إيجاد رتبة المئين المطلوب} = (\text{ترتيب المئين} / 100) * (\text{عدد البيانات} + 1)$$

$$\text{Rank } P_k = k(n+1)/100$$

$$\text{رتبة المئين} = \frac{k * (n+1)}{100}$$

* إيجاد المئين للبيانات المبوبة **Grouped Data**

إيجاد موقع أي مئين يتم تحديده كالتالي: $(i \sum f_i) / 100$

موقع المئين ثمانين P80 هو القيمة $(80 \sum f_i) / 100$

تحديد التكرار الصاعد المقابل للمئين المطلوب.

تحديد فئة المئين المطلوب: وهي أول فئة تقابل تكرار صاعد يساوي أو يزيد عن التكرار

الصاعد للمئين.

المئين ك = الحد الأدنى الفعلي لفئته + $(100/ك) * (\text{مجموع (ت) - تكرار صاعد للفئة السابقة}) * ط$

تكرار عادي لفئة المئين

إن قيمة المئين i تحسب عن طريق الصيغة التالية:

$$P_i = L + \frac{[(i \sum f_i) / 100] - f_1}{f_2 - f_1} * H$$

3-1 مقاييس التشتت Measure Dispersion or Variation

تستخدم مقاييس التشتت لقياس انتشار قيم المشاهدات حول نقطة التركيز وهي الوسط الحسابي، وإن المقصود بالتشتت أو الاختلاف هو التباعد الموجود بين قيم المشاهدات التابعة لمتغير ما عن وسطها الحسابي.

إذا كان لدينا مجموعتان من البيانات لها نفس الوسط الحسابي فمن الممكن أن تكونا مختلفتين في انتشارهما حول الوسط الحسابي.

كلما كبرت قيم مقاييس التشتت دل ذلك على درجة كبيرة من الاختلاف بين قيم البيانات، وكلما صغرت قيم مقاييس التشتت دل ذلك على درجة قليلة من الاختلاف بين قيم البيانات.

بذلك هذه المقاييس تعطي فكرة عن مدى تجانس أو اختلاف البيانات عن مركزها ويدل ذلك على درجة انتشارها.

وهناك عدة مقاييس للتشتت وهي:

1. المدى Range (R)

2. الانحراف المتوسط Mean Deviation (MD)

3. التباين Variance (σ^2)

4. الانحراف المعياري Standard Deviation (σ)

5. معامل الاختلاف Coefficient Variation (CV)

1- المدى Range (R)

يعرف المدى لمجموعة من البيانات على أنه الفرق بين أكبر قيمة وأصغر قيمة لتلك المجموعة، ويرمز له بالرمز R.

* المدى للبيانات غير المبوبة Ungrouped Data

المدى = أكبر قيمة - أصغر قيمة Range = Highest score - Lowest score

* المدى للبيانات المبوبة Grouped Data

بما أن أصغر قيمة وأكبر قيمة مجهولة في حالة المعطيات المبوبة فإن قيمة المدى التقديرية

هي:

المدى للتوزيع = مركز الفئة الأخيرة - مركز الفئة الأولى
= الحد الأعلى للفئة الأخيرة - الحد الأدنى للفئة الأولى
= الحد الأعلى الفعلي للفئة الأخيرة - الحد الأدنى الفعلي للفئة الأولى

* عيوب المدى Range's Deficiencies

1. يتأثر بالقيم الشاذة والمتطرفة.

2. لا يستخدم في حالة المعطيات المبوبة التي تتضمن فئات مفتوحة.

* مزايا المدى Range's Characteristics

1. سهولة الفهم.
2. سهولة حسابه.
3. كثرة استخدامه في الأوساط العامة.

2- الانحراف المتوسط Mean Deviation (MD)

الوسط الحسابي للقيمة المطلقة لانحراف تلك القيم عن وسطها الحسابي.
معدل انحراف المشاهدات في التوزيع عن وسطه الحسابي.

* الانحراف المتوسط للبيانات غير المبوبة Ungrouped Data

$$A.M = \frac{\sum |x_i - \bar{x}|}{n}$$

$$MD = \frac{\sum |x_i - \bar{x}|}{n}$$

حيث: $M.D$ ، أ.م : الانحراف المتوسط.

x_i ، س: المشاهدات. \bar{x} ، ق: الوسط الحسابي.

* الانحراف المتوسط للبيانات المبوبة Grouped Data

بافتراض أن كافة القيم الواقعة ضمن الفئات هي عند مركز هذه الفئات فإن الانحراف المتوسط عبارة عن مجموع انحرافات قيم مراكز هذه الفئات عن الوسط الحسابي مضروبة بتكرارها ثم قسمة الناتج على مجموع التكرارات.

$$A.M = \frac{\sum |x_i - \bar{x}| * f_i}{\sum f_i}$$

$$MD = \frac{\sum D_i * f_i}{\sum f_i}$$

حيث:

$|s - \bar{s}|$ ، D_i : قيم الانحرافات المطلقة عن الوسط الحسابي $|x_i - \bar{x}|$
 X_i : مراكز الفئات.
 F_i : التكرار.

* عيوب الانحراف المتوسط Mean Deviation Deficiencies

1. نادر الاستخدام بسبب كون عملية احتسابه تعتمد على القيم المطلقة والتي تحمل الإشارة.

2. عدم إمكانية استخدامه مع الجداول التكرارية ذات الفئات المفتوحة.

* مزايا الانحراف المتوسط Mean Deviation Characteristics

1. في احتسابه يتم شمول كل القيم المطلوب تقدير قيمة تشتتها.

3- التباين (σ^2) Varlance

هو الوسط الحسابي لمربعات انحرافات القيم X_i عن وسطها الحسابي \bar{X} .
 القيمة المرتفعة للتباين تعني أن الأشياء متباينة، متباعدة، متناثرة، غير متجانسة.
 القيمة المنخفضة للتباين تعني أن الأشياء غير متباينة، متقاربة، متجانسة.

التباين للبيانات غير المبوبة Ungrouped Data

$$\sigma^2 = \frac{\sum (s - \bar{s})^2}{n}$$

$$\sigma^2 = \frac{\sum (s^2 - \bar{s}^2)}{n}$$

$$\sigma^2 = \frac{\sum (x_i - \bar{X})^2}{n}$$

$$\sigma^2 = \frac{\sum x_i^2 - n\bar{X}^2}{n}$$

$$\sigma^2 = \frac{\sum x_i^2 - [(\sum x)^2 / n]}{n}$$

* التباين للبيانات المبوبة Grouped Data

$$\sigma^2 = \frac{\sum (x_i - \bar{X})^2 * f_i}{\sum f_i}$$

$$\sigma^2 = \frac{\sum X_i^2 * f_i - n * \bar{X}^2}{\sum f_i}$$

4- الانحراف المعياري (σ) Standard Deviation

الانحراف المعياري من أكثر مقاييس التشتت شيوعاً وأهمية، وهو الجذر التربيعي لمتوسط مربعات الانحرافات القيم عن متوسطها الحسابي، وإن الانحراف المعياري هو الجذر التربيعي للتباين.

حساب الانحراف المعياري:

- إيجاد المتوسط الحسابي للمشاهدات.
- إيجاد انحرافات القيم المختلفة عن المتوسط الحسابي ومن ثم تربيعها، ثم جمعها.
- إيجاد متوسط مربعات انحرافات القيم عن متوسطها الحسابي.
- إيجاد الجذر التربيعي لمتوسط مربعات انحرافات القيم عن متوسطها الحسابي.

القانون الأساسي للانحراف المعياري:

$$\sigma = \sqrt{\sum (X - \bar{X})^2 / n} = \sqrt{\sigma^2}$$

الانحراف المعياري = الجذر التربيعي للتباين

$$\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}} = \sqrt{\sigma^2}$$

الشكل العملي لقانون الانحراف المعياري

تعتمد هذه الصيغة على القيم الأصلية، ولا يدخل المتوسط أو الانحرافات عن الوسط الحسابي في حسابها بشكل مباشر، وتكون المعادلة على الشكل التالي:

$$\sigma = \sqrt{\left\{ \frac{\sum x^2}{n} \right\} - \left\{ \frac{\sum x}{n} \right\}^2}$$

$$\sigma = \sqrt{(\sum x_i^2 - [(\sum x)^2 / n]) / n}$$

حيث أن: X: المشاهدات. n: عدد المشاهدات.

* خصائص الانحراف المعياري Standard Deviation Characteristics

- 1- يعتبر من أكثر مقاييس التشتت استخداماً وأهمية.
- 2- يتأثر بالتحويلات الخطية.
- 3- يعتمد في حسابه على جميع المشاهدات.
- 4- في التوزيعات القريبة من التوزيع الطبيعي نلاحظ أن:
 - 68.27% من البيانات تقع في مدى انحراف معياري واحد عن الوسط.
 - 95.45% من البيانات تقع في مدى انحرافين معياريين عن الوسط.
 - 99.73% من البيانات تقع في مدى ثلاثة انحرافات معيارية عن الوسط.

* الانحراف المعياري للبيانات المبوبة Grouped Data

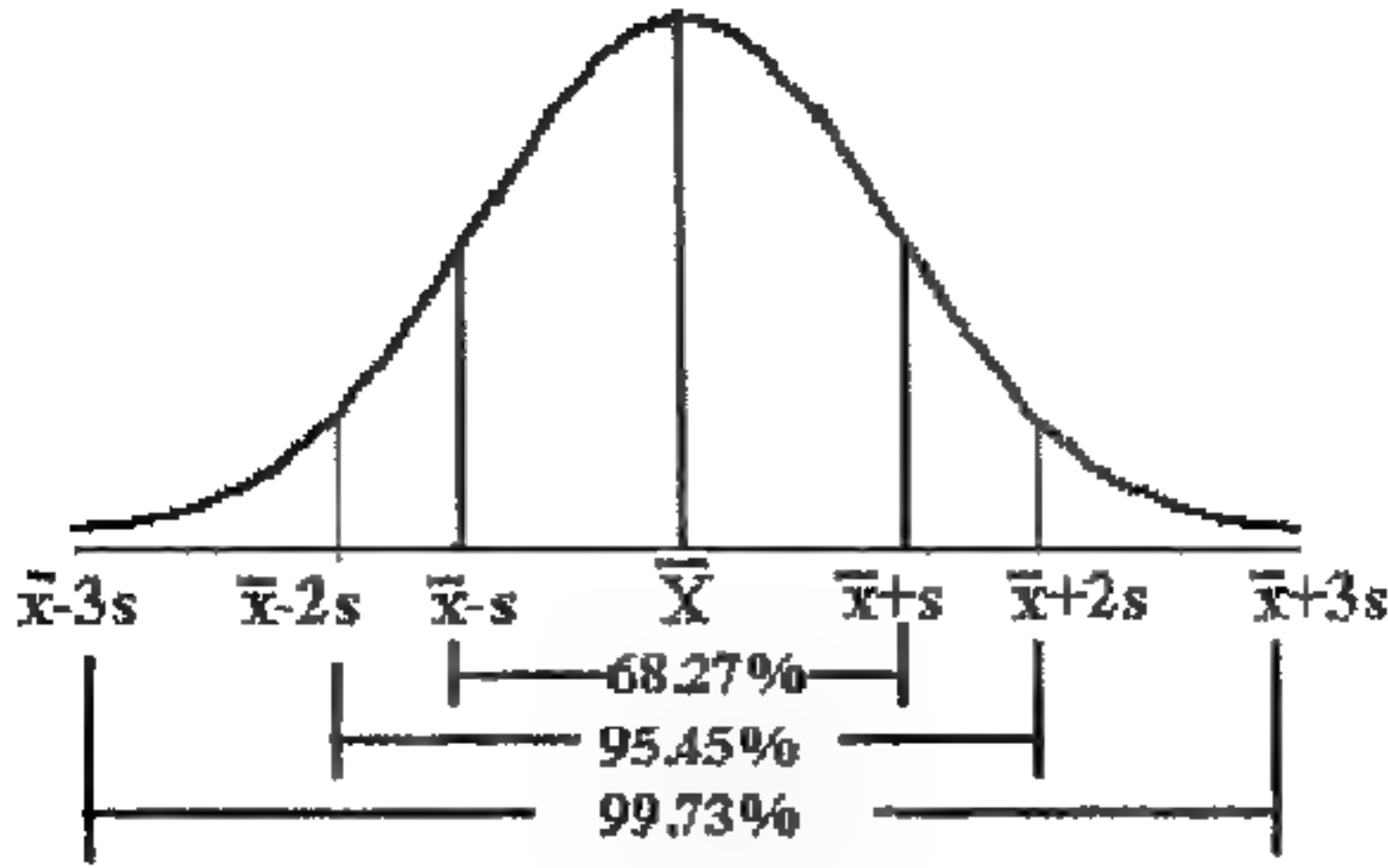
$$\sigma = \sqrt{\frac{\sum f_i x_i^2 - \frac{(\sum f_i x_i)^2}{n}}{n-1}}$$

حيث:

X_i : مركز الفئة. f_i : التكرار. n : عدد المشاهدات.

* تفسير الانحراف المعياري Standard Deviation Explanation

يستخدم الانحراف المعياري بشكل عام كأحد مقاييس التشتت، وذلك لمعرفة عدد المشاهدات ونسبتها والتي تقع على بعد معين من الوسط الحسابي \bar{X} حيث أنه وباستخدام الصيغة التجريبية (في التوزيعات المتماثلة والطبيعية) نلاحظ أن نسب المشاهدات تكون كما هو موضح بالشكل (1).



شكل (1): الصيغة التجريبية للتوزيع المتماثل.

في التوزيعات الطبيعية نتوقع أن 68.27% من القيم تقع على بعد انحراف معياري حول الوسط، وأن 95.45% من القيم تقع على بعد انحرافين معياريين حول الوسط، وأن 99.73% من القيم تقع على بعد ثلاث انحرافات معيارية حول الوسط، أما في التوزيعات غير المتماثلة نتوقع الحصول على نسب مختلفة عن تلك.

ولتوضيح عملية استخدام الانحراف المعياري لتفسير عينة ما، يجب إيجاد الفترات $\bar{X} \pm s$ ، $\bar{X} \pm 2s$ ، $\bar{X} \pm 3s$ ثم تحديد عدد البيانات ثم نسبتها، وبعد ذلك مقارنتها مع النسب المعيارية [68.27%, 95.45%, 99.73%] لتحديد ما إذا كان توزيع المشاهدات طبيعي والتشتت معنبد، أما لمقارنة عينتين فيمكن مقارنة النسب ببعضها والحكم بعد ذلك أن كان تشتت إحدى العينتين أكبر أو أقل من العينة الأخرى.

5- معامل الاختلاف (CV) Coefficient of Variation

إن معامل الاختلاف يفضل عن الانحراف المعياري وغيره من مقاييس التشتت لمقارنة تشتت البيانات بين عدة مجموعات من البيانات.

إن معامل الاختلاف يعطي نسبة الانحراف المعياري إلى الوسط الحسابي، وبما أن معامل الاختلاف هو مقياس لقياس التغير النسبي على شكل نسبة مئوية، لذلك معامل التغير يمكن استخدامه لمقارنة التشتت داخل عدة مجموعات من البيانات حتى وإن كانت وحدات القياس لهذه المجموعات مختلفة.

ويحسب من خلال المقارنة بين الوسط الحسابي والانحراف المعياري للبيانات حسب :

$$CV = \frac{S_x}{\bar{X}} * 100\% \quad \text{or} \quad CV = \frac{\sigma}{\mu} * 100\%$$

معامل الاختلاف = (الانحراف المعياري/الوسط الحسابي) * 100%

كلما كان معامل التغير كبير يعني ذلك وجود تباين واختلاف بين المشاهدات.

4-1 الارتباط Correlation والانحدار Regression

إن نظرية الارتباط تظهر قوة العلاقة بين متغيرين مع إمكانية تحديد نوع وقوة العلاقة بين الظواهر، كالعلاقة بين مستوى التعليم والأداء، والعلاقة بين معدل الثانوية العامة ومعدل الجامعة، والعلاقة بين المستوى الاقتصادي والتحصيل.

إن الهدف من تحليل الارتباط Correlation هو معرفة وجود علاقة بين متغيرين أو مجموعة من المتغيرات المستقلة Independent Variables (X_i) مع المتغير التابع Variable Dependent (Y_i) مسن عدم وجودها، وهناك عدة مقاييس لتحديد درجة العلاقة والارتباط بين المتغيرات.

الارتباط Correlation: علاقة بين متغيرين لمعرفة ما إذا كان تغير أحدهما مرتبطاً بتغير الآخر.

أنواع الارتباط Types of Correlation

* من حيث القوة Based on Strength

1. ارتباط تام Complete Correlation: يتحدد متغير كلياً عن طريق متغير آخر.

مثال) العلاقة بين مساحة الدائرة ونصف قطرها.

2. ارتباط غير تام Partial Correlation: يتأثر متغير معين بمتغير آخر.

مثال) حجم الإنفاق ودخل الأسرة.

* من حيث عدد المتغيرات Based on Number of Variables

1. ارتباط بسيط Simple Correlation: r ارتباط بين متغيرين كميين فقط.

مثال) العلاقة بين عمر الأب وعمر الطالب.

2. ارتباط متعدد **Multiple Correlation: R** يدرس العلاقة بين أكثر من متغيرين، تكون العلاقة بين المتغير التابع Y وعدة متغيرات مستقلة X_i .
(مثال) العلاقة بين حجم المبيعات وعدد الأسواق وعدد السكان.

* قياس الارتباط **Measures of Correlation**

أ- معامل ارتباط بيرسون الخطي **Person Linear Correlation Coefficient**

معامل ارتباط بيرسون يقيس قوة واتجاه العلاقة الخطية فقط بين متغيرين كميين.

ب- معامل ارتباط سبيرمان للترتيب **Spearman Correlation Coefficient**

يستخدم معامل ارتباط سبيرمان ومعامل كندال تاو لقياس قوة الارتباط بين متغيرين ترتيبيين **Ordinal**.

الانحدار ومفهومه **Regression Concept**

في دراسة العلاقة بين ظاهرتين أو أكثر إذا كان الهدف تحديد نوع وقوة العلاقة فإننا ندرس الارتباط **Correlation**، أما إذا كان الهدف دراسة العلاقة من خلال التمثيل البياني بأفضل علاقة اقتران ممكنة بالشكل $Y_i = f(x)$ فإننا ندرس الانحدار **Regression** ويسمى المستقيم أو المنحنى الذي يمثل هذه الاقتران بمستقيم أو منحنى الانحدار، وهو من الأساليب الإحصائية المستخدمة لتحديد التأثيرات بين المتغيرات المستقلة والمتغير التابع عن طريق معادلة الانحدار للتنبؤ بقيمة المتغير التابع بدلالة المتغيرات المستقلة، فإذا كان عدد المتغيرات المستقلة واحد فيسمى انحدار خطي بسيط **Simple Linear Regression**، أما إذا كان عدد المتغيرات المستقلة أكثر من واحد فيسمى انحدار متعدد **Multiple Regression**

الانحدار **Regression**:

للانحدار عدة تعريفات منها:

■ الميل أو الانحدار نحو الوسط.

■ إيجاد النموذج الذي يمثل العلاقة السببية بين متغيرين أو أكثر.

■ يستخدم للتنبؤ بقيمة متغير عن طريق معرفة متغير آخر مرتبط به، مثل التنبؤ بالأرباح إذا عرفت المبيعات.

■ العلاقة بين المتغيرات من خلال بناء معادلة تستخدم للتقدير والتنبؤ بقيمة المتغير التابع Y بدلالة متغير أو متغيرات مستقلة X_i ، مثل العلاقة بين الدخل والطلب.

■ الانحدار الخطي Linear Regression: تكون معادلة التقدير عند عرضها بيانياً على شكل خط مستقيم.

التنبؤ:

تقدير بيانات غير معروفة مبنية على بيانات معروفة وذات صلة بالظاهرة.
الانحدار يفترض وجود علاقة خطية قوية.

تحليل الانحدار Regression Analysis:

الأساليب التي تستخدم في تقدير قيمة متغير عند معرفة قيم متغير آخر.
تحليل الانحدار الخطي الثنائي Bivariate Linear Regression: يستخدم لتمثيل العلاقة على شكل معادلة خطية للتنبؤ بقيمة متغير من خلال قيم متغير آخر، ويكون المتغير الأول كمياً ويسمى المتنبئ ويكون المتغير الثاني كمياً ويسمى المتنبأ به.

* أهداف تحليل الانحدار Regression Analysis Objectives

1. تحديد العلاقة بين المتغير التابع Y والمتغيرات المستقلة X_i
2. التنبؤ بقيمة المتغير التابع Y عن طريق المتغيرات المستقلة X_i
3. الاستنتاج حول المجتمع من خلال المعادلة التقديرية.
4. اختبار الفروق بين خط الانحدار التقديري وخط الانحدار الحقيقي.

* الانحدار الخطي البسيط Simple Linear Regression

هو البحث في العلاقة بين متغيرين فقط وهما المتغير التابع Y والمتغير المستقل X ، وإن معادلة الانحدار في المجتمع هي:

$$Y_i = \alpha + \beta X_i + \varepsilon_i$$

Y : المتغير التابع. X : المتغير المستقل. α : معامل ثابت. β : ميل الانحدار.

ε_i : الأشياء الأخرى التي تسبب الخطأ Disturbance term error

المتغير المستقل Independent Variable: المتغير الذي يستخدم في تقدير قيم المتغير الآخر (السبب).

المتغير التابع Dependent Variable: المتغير الذي نقيده (الآثر).

التباين Variance: وجود فروق بين المشاهدات، مسافات، إذا كانت كبيرة يكون

التباين كبير، وإذا كانت صغيرة يكون التباين صغير، وإذا انعدمت انعدم التباين.

إذا اعتمدت معادلة الانحدار على بيانات العينة نستعوض عن α بالرمز a وعن β بالرمز

b وتصبح

$$\hat{y}_i = a + bX_i + e_i$$

معادلة الانحدار في العينة هي:

$$e = (Y - \bar{Y})$$

$$\hat{y} = a + bX_i$$

Y : المتغير التابع. X : المتغير المستقل. a : ثابت. b : معامل انحدار.

إن خط الانحدار يمر من خلال نقطة الوسط (\bar{X}, \bar{Y}) ، لذلك فإن مجموعة الأخطاء دائماً

تساوي صفر $\sum e_i = 0$

* معادلة خط الانحدار Regression Equation

1- المعادلة التقديرية للانحدار الخطي البسيط Equation Estimation

$$Y_i = E(Y_i) + \varepsilon_i$$

$$Y_i = a + bX_i + e_i$$

e_i : الخطأ العشوائي.

2- طريقة المربعات الصغرى Least Squares Method

حتى نقدر ميل خط الانحدار غير المعلوم بواسطة طريقة المربعات الصغرى والتي تعتمد

على تقليل مجموع مربعات انحرافات القيم الحقيقية y_i عن القيم التقديرية \hat{y}_i حيث (b) تمثل الميل

وهو نسبة تغير قيمة المتغير التابع y إلى وحدة واحدة من المتغير المستقل x وان a تمثل معامل التقاطع والذي يعني مقدار قيمة y عندما تكون قيمة المتغير المستقل $x = 0$ صفرًا. وذلك بإتباع الخطوات التالية:

- معادلة خط الميل التقديري $\hat{y} = a + bX$ حيث \hat{y} هي تقدير $E(y)$, α , β
- انحراف كل قيمة حقيقية y_i عن القيمة التقديرية \hat{y}_i يكون بمقدار e_i وهي $e_i = y_i - \hat{y}_i$

$$b = \frac{\sum X_i Y_i - n \bar{X} \bar{Y}}{\sum X_i^2 - n \bar{X}^2}$$

$$a = \bar{Y} - b * \bar{X}$$

- ويمكن حساب معامل الانحدار b إذا علمنا بقيمة معامل الارتباط البسيط بين x و y

كالتالي:

$$b = \frac{S_y}{S_x} * r_{xy}$$

$$r_{xy} = \frac{S_{xy}}{S_x S_y}$$

- بتعويض قيم a , b نحصل على خط الانحدار التقديري كتقدير لخط انحدار المجتمع

$$E(Y) = \alpha + \beta X$$

أ- معادلة خط انحدار ص على س هي:

$$\text{ص} = \text{أ} + \text{ب س} \quad \text{التنبؤ بقيمة ص إذا علمت قيمة س}$$

$$\text{ب} = (\sigma_v / \sigma_s) * r$$

$$\text{ب} = \frac{(\sum \text{ص ر} - (\sum \text{ص} * \sum \text{ر}) / \text{ن})}{\sum \text{ر}^2 - (\sum \text{ر})^2 / \text{ن}}$$

$$\text{أ} = \bar{\text{ص}} - \text{ب} * \bar{\text{س}}$$

ب- معادلة خط الميول S على Y هي:

$$S = A + B Y$$

التي B قيمة S إذا علمت قيمة Y

$$B = \frac{\sigma_Y}{\sigma_X} r$$

$$B = \frac{(\sum XY - \frac{(\sum X)(\sum Y)}{n})}{(\sum X^2 - \frac{(\sum X)^2}{n})}$$

$$A = \bar{Y} - B \bar{X}$$

* خواص مقدرات طريقة المربعات الصغرى

■ خاصية عدم التحيز. ■ خاصية أقل تباين ممكن.

Multiple Linear Regression

* الانحدار الخطي المتعدد

الانحدار المتعدد (متغيرين مستقلين) Two Independent Variables

تحليل الانحدار ملائم لمعرفة كم من التباين يستطيع المتغيران معاً أن يفسرانه من التباين في المتغير التابع.

بمعرفة تباين المقدار الذي يشرحه المتغير الأول، ومعرفة تباين المقدار الذي يشرحه المتغير الثاني، نستطيع أن نكشف عن الأهمية النسبية لكلا المتغيرين.

الافتراضات للانحدار المتعدد

1. أن المتغيرات المستقلة مستقلة (معامل الارتباط بينها 0).

2. أن المتغيرات الداخلة على يمين إشارة = في المعادلة يمكن أن يكون بينها ترابطات

المتغيرات المستقلة (متنبئات) Predictors المتغير التابع (محك) Criterion

* معادلة الانحدار الخطي المتعدد

$$\hat{Y} = a + b_1 X_1 + b_2 X_2$$

a : المقطع الصادي وهو بعد النقطة التي يقطع عندها خط الانحدار.

b_1 : معامل الانحدار للمتغير الأول x_1 (وزن الانحدار)، تعني التغير في y الذي يقابل تغير x_1 وحدة واحدة مع بقاء الأشياء الأخرى ثابتة.

b_2 : معامل الانحدار للمتغير الثاني x_2 ، تعني التغير في y الذي يقابل تغير x_2 وحدة واحدة مع بقاء الأشياء الأخرى ثابتة.

اختيار الأساليب الإحصائية الوصفية المستخدمة لتغير واحد:

نوع المتغير	أساليب القياس المناسبة			
	نوعه مركزية	تشتت	المقاييس النسبية	أخرى
اسمي Nominal	المتوال	التكرار النسبي للقيمة المتوالية	التكرار النسبي (% للتكرارات)	-----
رتبي Ordinal	الوسيط	نصف المدى الربيعي	التكرار النسبي	-----
فئوي أو نسبي Interval Or Ratio	المتوسط للتوزيع الاعتدالي. الوسط والمتوسط للتوزيع ملتر	المدى المطلق. التباين الانحراف المعياري	التكرار النسبي مثل النسبة، المئين الارباعيات	معاملات لاشور والشفرطع.

5-1 استخدام برمجية SPSS في حل المسائل.

مثال (1-1)

إذا كانت الاستبانة التالية هي عبارة عن استبانة (10) طلاب عني سؤال بتدريج ليكرت الخماسي وهو (ان التعليم العالي في الاردن في تطور مستمر)، وعلامات الطلاب بمادتي الاحصاء والقياس، كما تحتوي بعض المعلومات عن الطلبة المستجيبين لهذه الاستبانة.

○ رقم الطالب :

الجنس ○ ذكر ○ انثى

المستوى ○ اولى ○ ثانية ○ ثالثة ○ رابعة

الرقم	الجنس	المستوى	علامة الاحصاء	علامة القياس	(ان التعليم العالي في الاردن في تطور مستمر) درجة الموافقة			
					موافق بشدة	موافق	محايد	معارض بشدة
-1	1	1	65	90	-			
-2	2	3	70	85	-			
-3	1	2	75	83	-			
-4	1	4	80	95	-			
-5	2	3	60	90			-	
-6	1	2	90	70				-
-7	1	3	80	75			-	
-8	2	2	65	60	-			
-9	2	2	80	80				-
-10	1	1	55	50				-

اجب عن الاسئلة التالية:

- 1- اعمل على ترميز المتغيرات الواردة بالسؤال اعلاه.
- 2- ادخل البيانات الى البرنامج.
- 3- احسب مقاييس الترتبة المركزية.
- 4- احسب مقاييس التشتت.
- 5- ما نسبة الذكور والإناث في عينة الدراسة؟
- 6- ما عدد أفراد العينة في كل مستوى؟
- 7- جد معامل الارتباط بين الجنس والمستوى وهل هو دال على مستوى ($\alpha=0.05$).
- 8- جد معامل الارتباط بين علامة الاحصاء وعلامة القياس وهل هو دال على مستوى ($\alpha=0.05$).
- 9- هل هناك علاقة بين مستوى الطالب ودرجة موافقته وهل هو دال على مستوى ($\alpha=0.05$) ولصالح من؟

10- اعمل على حفظ الملف تحت اسم Q1

الحل:

1- اعمل على ترميز المتغيرات الواردة بالسؤال اعلاه.

عملية تحويل إجابات كل سؤال إلى أرقام أو حروف يسهل إدخالها إلى الحاسوب.

- متغير الجنس Sex وهو (ذكر، أنثى)، حيث يعطى الرقم 1 للذكور والرقم 2

للإناث.

- المستوى الدراسي:

○ سنة أولى. 1 سنة ثانية. 2

○ سنة ثالثة. 3 سنة رابعة. 4

- أسئلة الاستبانة مصممة على أساس مقياس "ليكرت" الخماسي كما يلي:

○ موافق بشدة. 5

○ موافق. 4

○ محايد. 3

○ معارض. 2

○ معارض بشدة. 1

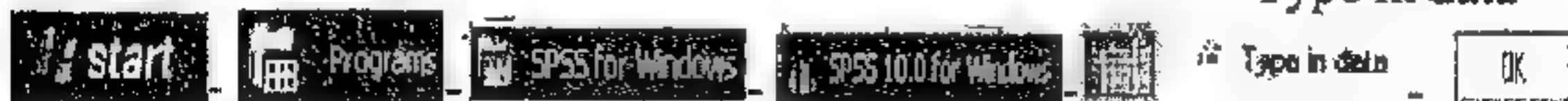
لعمل جدول ترميز للمتغيرات الواردة في مثال 1-1 ، بواسطة استخدام برنامج SPSS

يتم كما يلي:

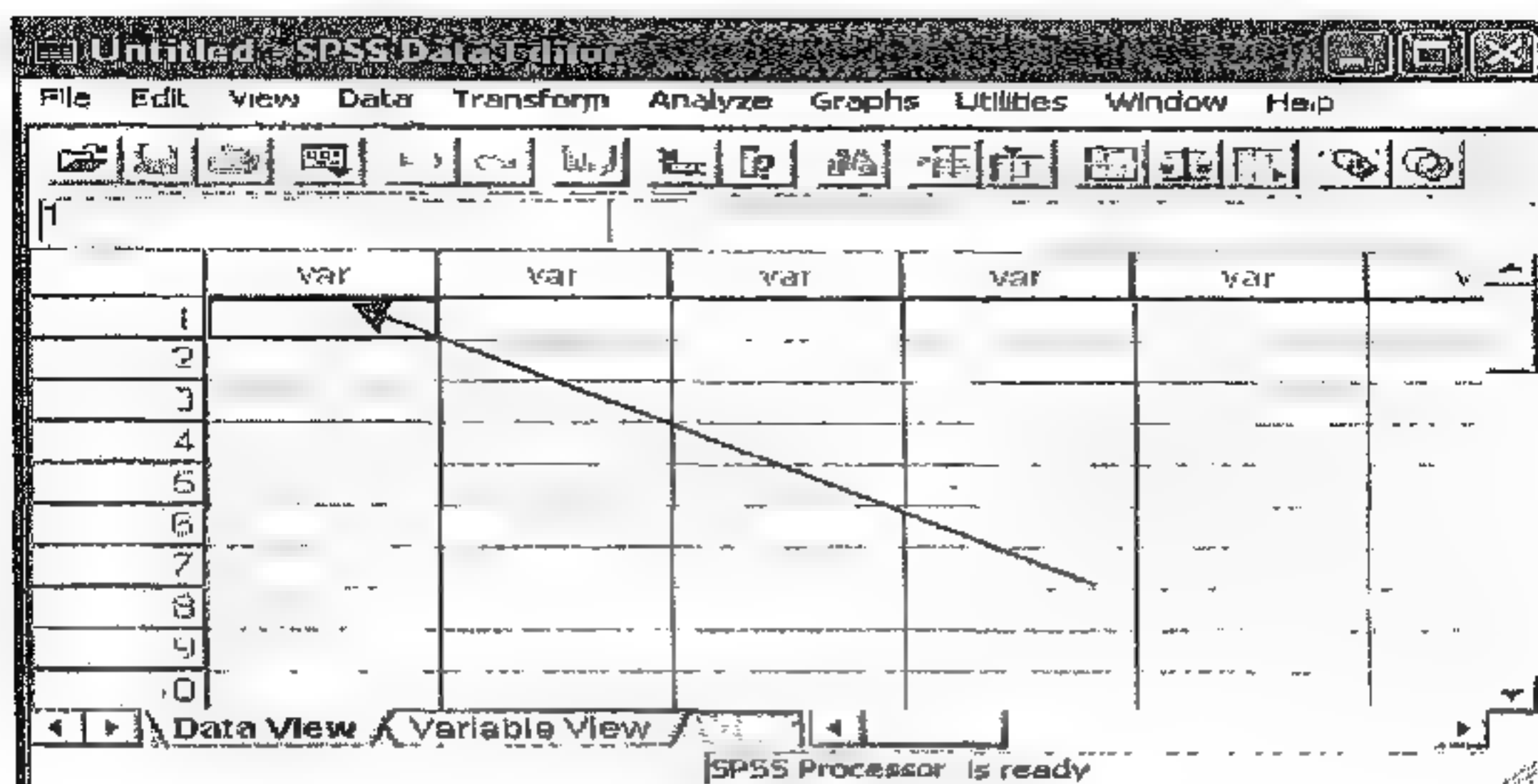
* تشغيل نظام SPSS

Start – Programs – SPSS for Windows – SPSS10.0for Windows –

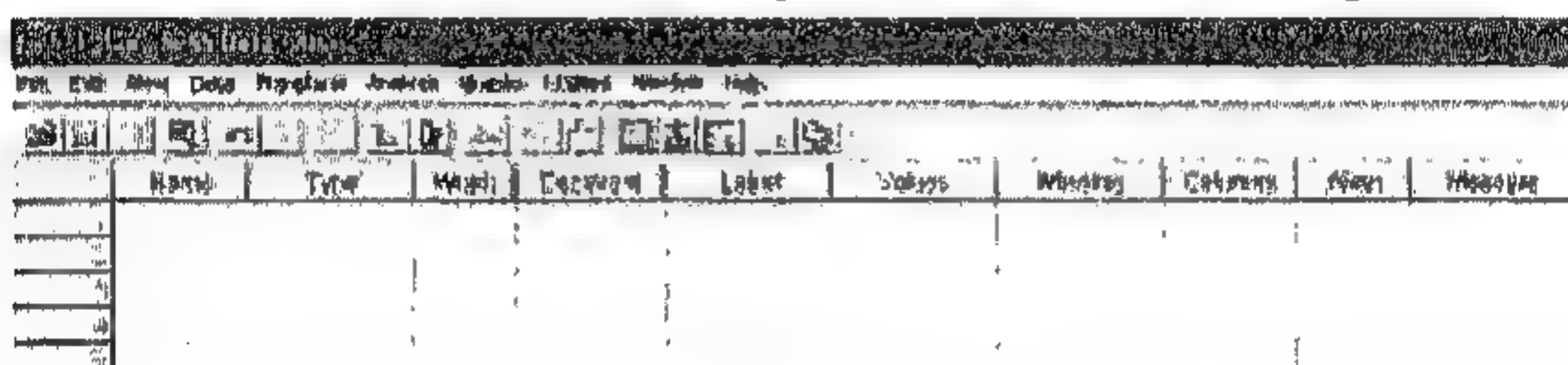
Type in data – Ok



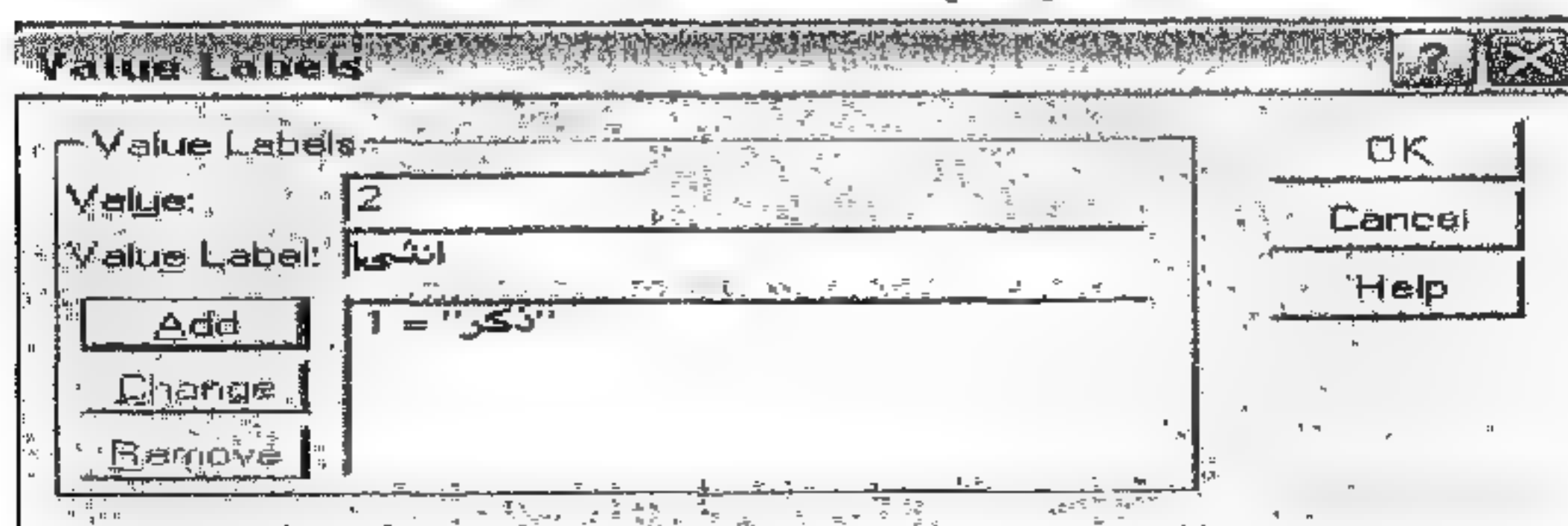
تظهر لديك الشاشة المبينة أدناه:



الخطوات المتبعة لتعريف المتغيرات: لنفرض أننا نريد تعريف المتغيرات الواردة في مثال 1-1
انقر على Variable View الموجودة على شريط الحالة فتظهر الشاشة أدناه:



نعيء الخانات أعلاه وهي Name ، Type ، ... ، Measure بالشكل التالي:
مثلا لادخال متغير الجنس نعيء الخانات أعلاه وهي Name اسم المتغير ونكتب الجنس،
ثم نوع المتغير Type ونختار Numeric أي رقمي، ... ، وعند وصولنا إلى خانة Values وهي
القيم المحتملة للمتغير نعيءها بالشكل التالي:



وبعد الانتهاء من تعريف المتغيرات تظهر الشاشة كما هو مبين أدناه:

Name	Type	Width	Decimals	Label	Values	Format	Columns	Align	Measure
الرقم	Numeric	10	0		Name	None	1	Center	Scale
الجنس	Numeric	10	0		Name	None	2	Center	Ordinal
المستوى	Numeric	10	0		Name	None	3	Center	Ordinal
ساعات	Numeric	10	0		Name	None	4	Center	Ordinal
الدرجة	Numeric	10	0		Name	None	5	Center	Scale
الوقت	Numeric	10	0		Name	None	6	Center	Scale

انقر على Data View الموجودة على شريط الحالة فتظهر الشاشة أدناه:

رقم	الجنس	المستوى	ساعات	الدرجة	الوقت
1					
2					
3					
4					
5					

2- ادخل البيانات الى البرنامج.

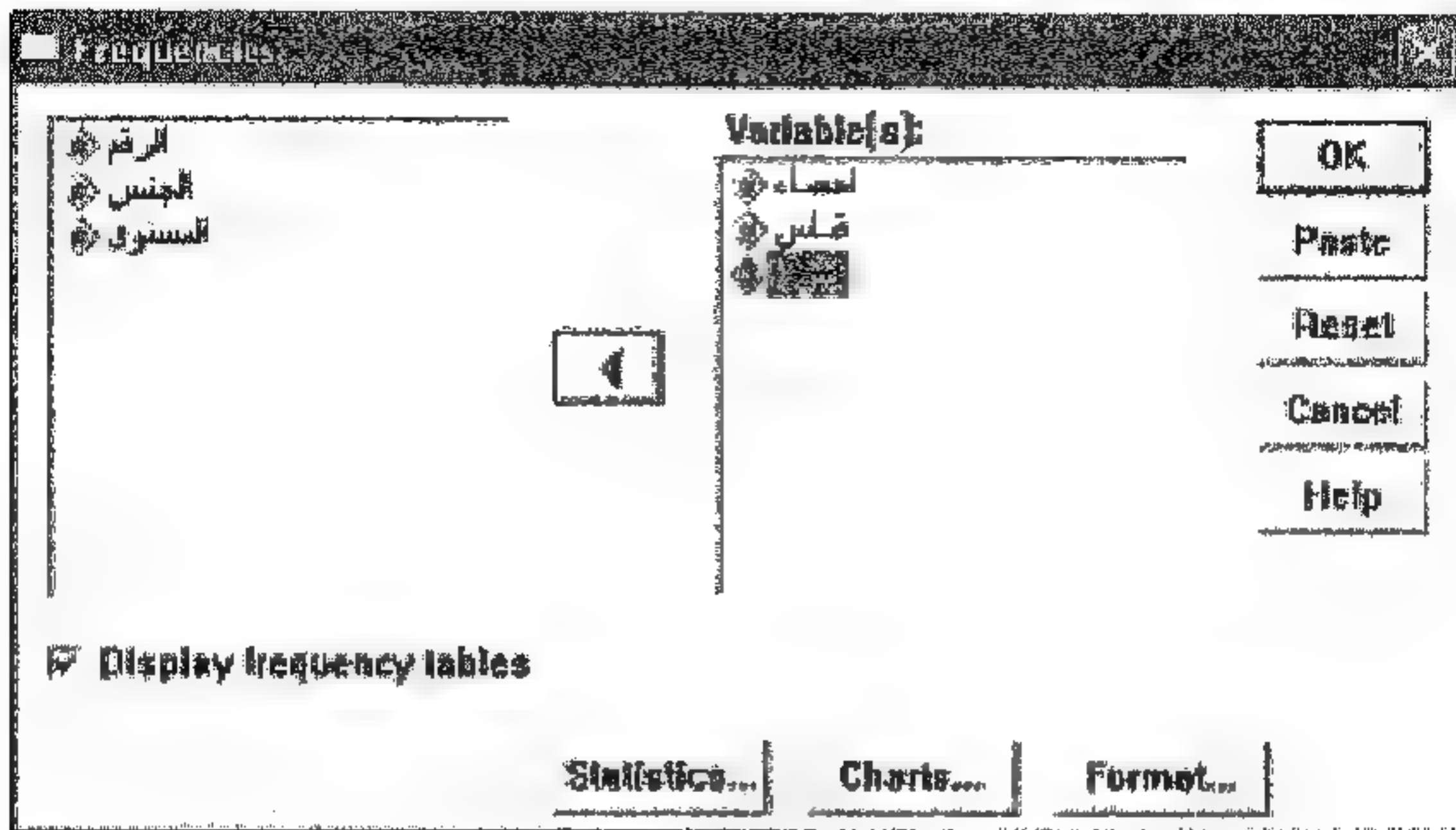
إدخال البيانات Input Data

الرقم	الجنس	المستوى	ساعات	الدرجة	الوقت
1	1	1	4	65	90
2	2	3	4	70	85
3	1	2	4	75	83
4	1	4	5	80	95
5	2	3	3	60	90
6	1	2	1	90	70
7	1	3	3	80	75
8	2	2	5	65	60
9	2	2	2	80	80
10	1	1	1	55	50

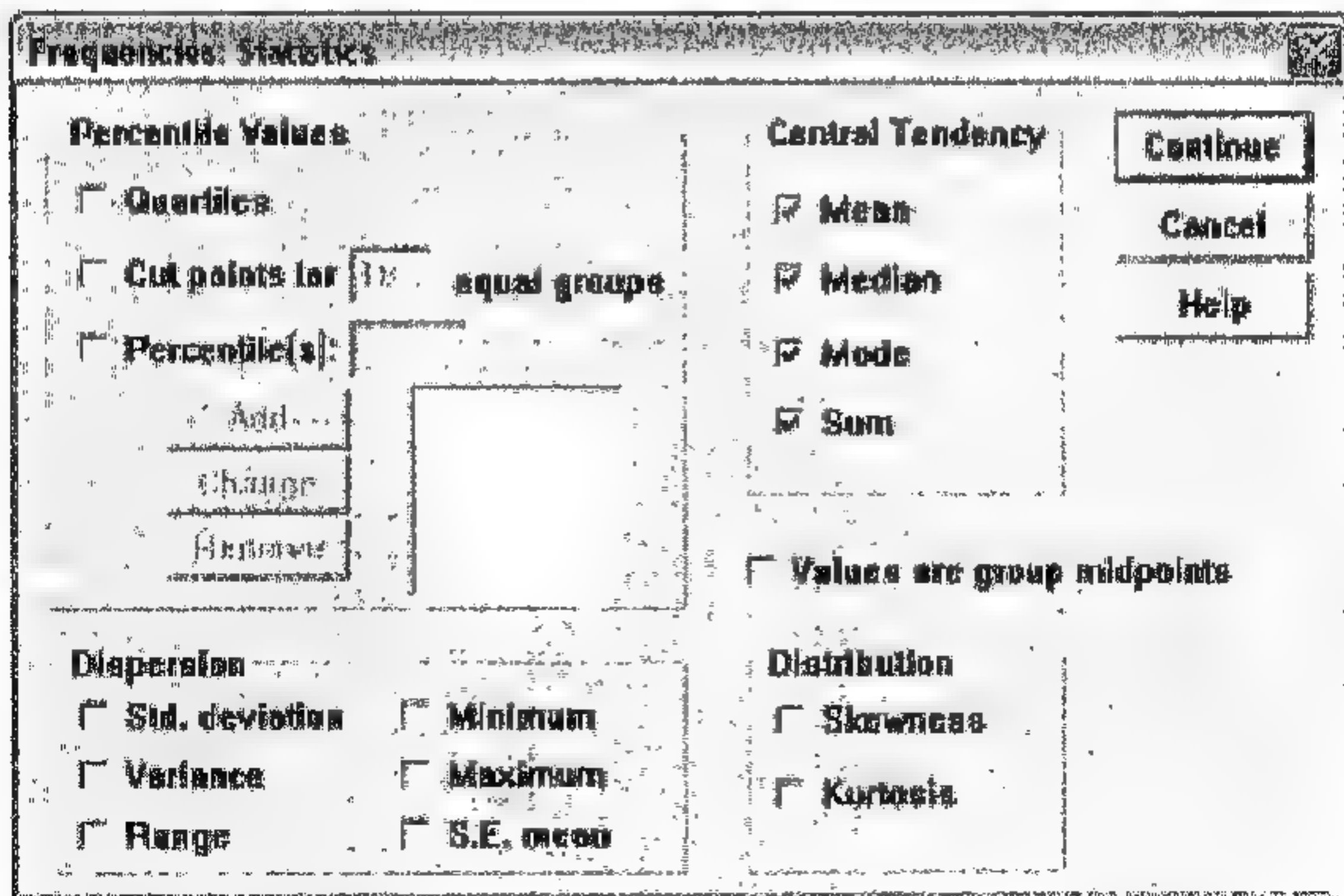
بعد الانتهاء من ادخال البيانات.

3- احسب مقاييس التزعة المركزية.

لعمل إحصاءات وصفية... Analyze Descriptive Statistics – Frequencies...

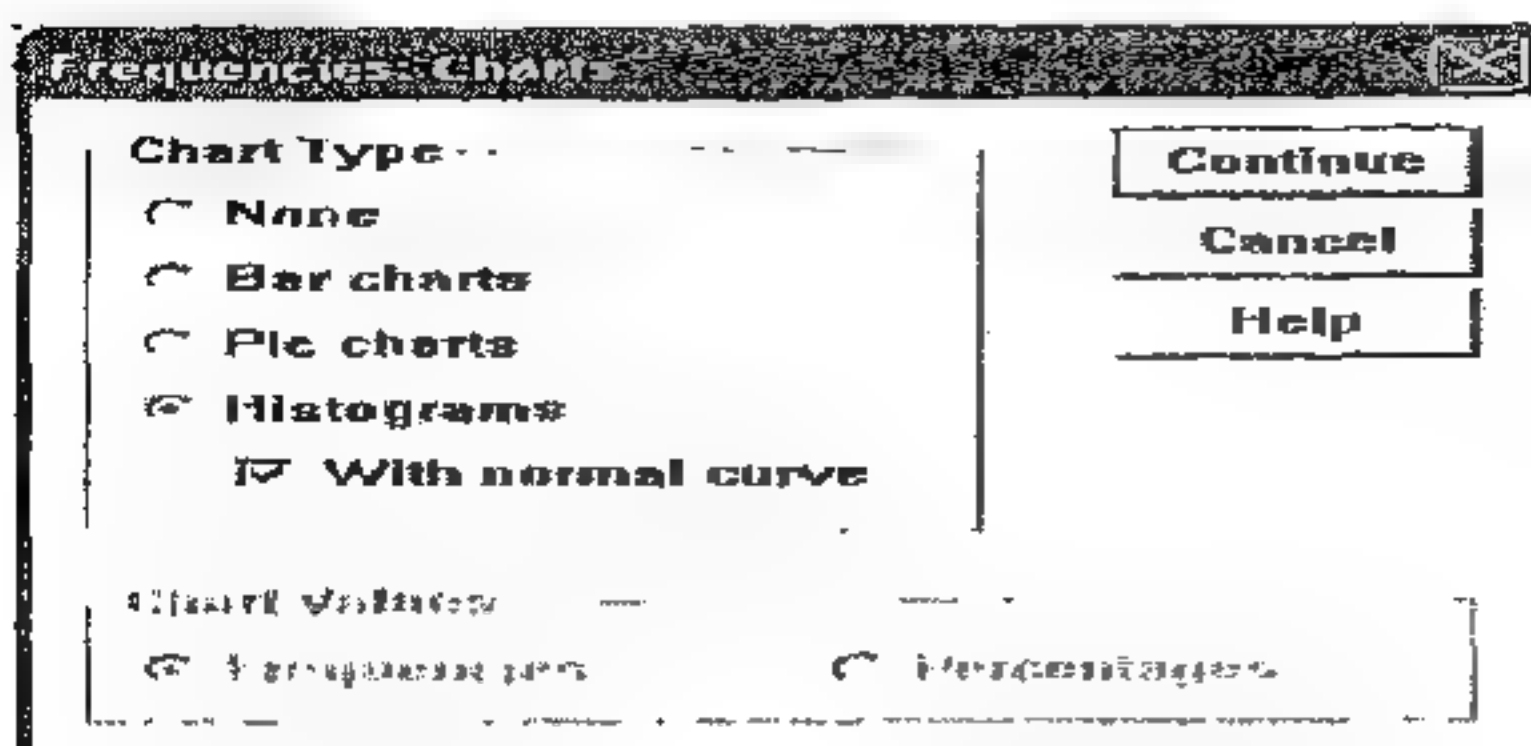


نعمل على وضع المتغيرات احصاء، قياس، س 1 في قائمة Variable(s) ثم نضغط زر Statistics فتظهر الشاشة أدناه:



نحدد مقاييس التزعة المركزية Central Tendency ثم نضغط زر Continue

نضغط زر Charts فتظهر الشاشة أدناه:



نحدد Histograms ونحدد with normal curve ثم نضغط زر Continue

ثم نضغط زر ok فتظهر شاشة المخرجات أدناه:

Statistics

		لصحاء	قيلين	س1
N	Valid	10	10	10
	Missing	0	0	0
Mean		72.00	77.80	3.20
Median		72.50	81.50	3.50
Mode		80	90	4
Sum		720	778	32

Frequency Table

لصحاء

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	68	1	10.0	10.0	10.0
	80	1	10.0	10.0	20.0
	85	2	20.0	20.0	40.0
	70	1	10.0	10.0	50.0
	75	1	10.0	10.0	60.0
	80	3	30.0	30.0	90.0
	90	1	10.0	10.0	100.0
Total		10	100.0	100.0	

قيلين

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	60	1	10.0	10.0	10.0
	80	1	10.0	10.0	20.0
	70	1	10.0	10.0	30.0
	75	1	10.0	10.0	40.0
	80	1	10.0	10.0	50.0
	83	1	10.0	10.0	60.0
	85	1	10.0	10.0	70.0
	90	2	20.0	20.0	90.0
	95	1	10.0	10.0	100.0
Total		10	100.0	100.0	

س1

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	شماره 1	2	20.0	20.0	20.0
	شماره 2	1	10.0	10.0	30.0
	شماره 3	2	20.0	20.0	50.0
	شماره 4	3	30.0	30.0	80.0
	شماره 5	2	20.0	20.0	100.0
Total		10	100.0	100.0	

4- احسب مقاييس التشتت.

لعمل إحصاءات وصفية... Analyze – Descriptive Statistics – Frequencies...

نعمل على وضع المتغيرات إحصاء، قياس، س1 في قائمة Variable(s) ثم نضغط

زر Statistics ونحدد مقاييس التشتت Dispersion ثم نضغط زر Continue ثم OK

Statistics

	Valid	Missing	الحصاة	القياس	س1
N	10	0	10	10	10
Std. Error of Mean	3.432	4.516			.457
Std. Deviation	10.853	14.381			1.476
Variance	117.778	203.966			2.178
Range	35	45			4
Minimum	66	60			1
Maximum	90	95			5

احصاء

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 66	1	10.0	10.0	10.0
60	1	10.0	10.0	20.0
70	2	20.0	20.0	40.0
75	1	10.0	10.0	50.0
76	1	10.0	10.0	60.0
83	3	30.0	30.0	90.0
90	1	10.0	10.0	100.0
Total	10	100.0	100.0	

قياس

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 60	1	10.0	10.0	10.0
60	1	10.0	10.0	20.0
70	1	10.0	10.0	30.0
75	1	10.0	10.0	40.0
80	1	10.0	10.0	50.0
83	1	10.0	10.0	60.0
85	1	10.0	10.0	70.0
90	2	20.0	20.0	90.0
95	1	10.0	10.0	100.0
Total	10	100.0	100.0	

س1

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid مناور صندل	2	20.0	20.0	20.0
مناور صندل	1	10.0	10.0	30.0
معالج	2	20.0	20.0	50.0
مواقي	3	30.0	30.0	80.0
مواقي	2	20.0	20.0	100.0
Total	10	100.0	100.0	

5- ما نسبة الذكور والإناث في عينة الدراسة؟

لعمل إحصاءات وصفية... Analyze – Descriptive Statistics - Frequencies

نعمل على وضع المتغير (الجنس) في قائمة Variable(s): ثم نضغط ثم OK

الجنس

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid ذكر	6	60.0	60.0	60.0
انثى	4	40.0	40.0	100.0
Total	10	100.0	100.0	

6- ما عدد أفراد العينة في كل مستوى؟

لعمل إحصاءات وصفية... Analyze – Descriptive Statistics – Frequencies

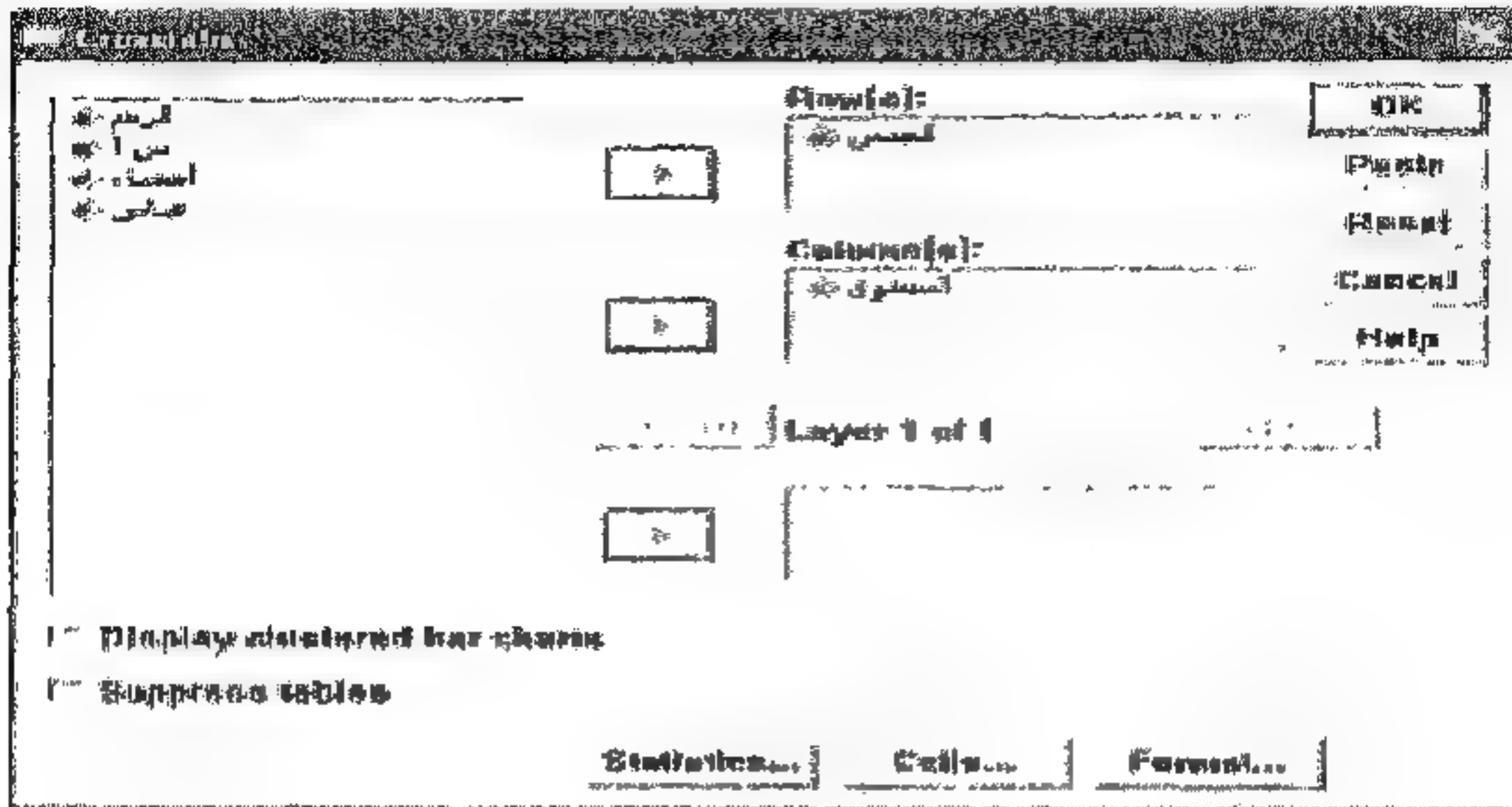
نعمل على وضع المتغير (المستوى) في قائمة Variable(s): ثم نضغط ثم OK

المستوى

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid سنفأولى	2	20.0	20.0	20.0
سنفأانية	4	40.0	40.0	60.0
سنفأانة	3	30.0	30.0	90.0
سنفأابعة	1	10.0	10.0	100.0
Total	10	100.0	100.0	

7- جد معامل الارتباط بين الجنس والمستوى وهل هو دال على مستوى ($\alpha=0.05$).

لعمل إحصاءات وصفية... Analyze – Descriptive Statistics – Crosstabs...



ثم نضغط ثم OK

Crosstabs

Case Processing Summary

	Crosstab					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
الجنس * المستوى	10	100.0%	0	0%	10	100.0%

Crosstab: الجنس * المستوى

Count	المستوى				Total
	متدبر	متدبر 2	متدبر 3	متدبر 4	
الذكور	3	2	1	1	6
الإناث	3	2	2	1	4
Total	6	4	3	2	10

Symptotic Measures

		Value	Asymptotic Sig.
Pearson Chi-Square	Pearson Chi-Square	.553	.383
	Continuity Correction ^a	.553	.383
N of Valid Cases		10	

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

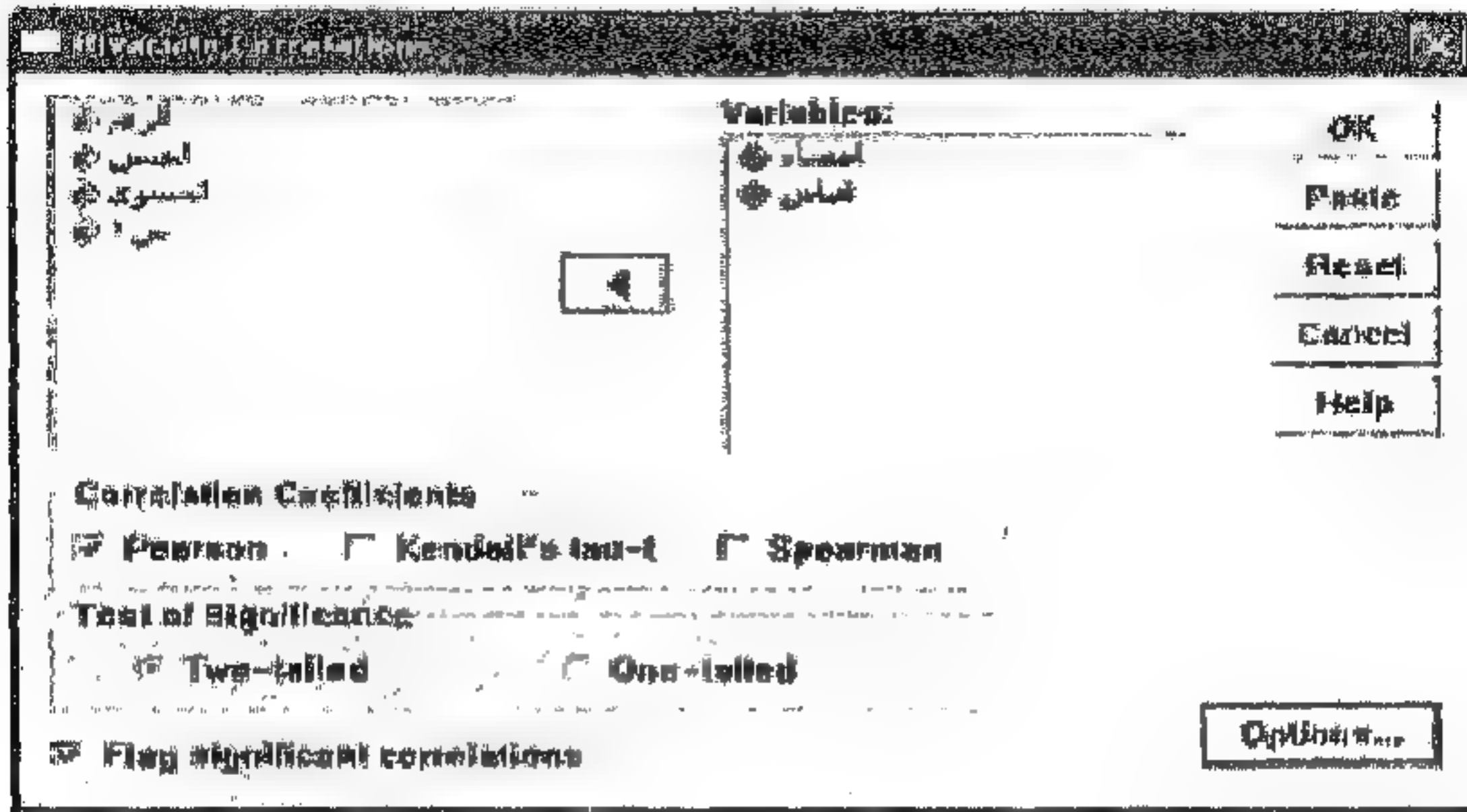
معامل الارتباط فاي بين الجنس والمستوى = 0.553

وهو غير دال إحصائياً لأن $\text{Sig}=0.383$ وهي أكبر من مستوى الدلالة ($\alpha=0.05$)

8- جند معامل الارتباط بين علامة الاحصاء وعلامة القياس وهل هو دال على مستوى $(\alpha=0.05)$.

Analyze – Correlate – Bivariate...

لحساب معامل الارتباط



ثم نضغط OK

→ Correlations

Descriptive Statistics

	Mean	Std. Deviation	N
الاحصاء	72.00	10.853	10
القياس	77.80	14.281	10

Correlations

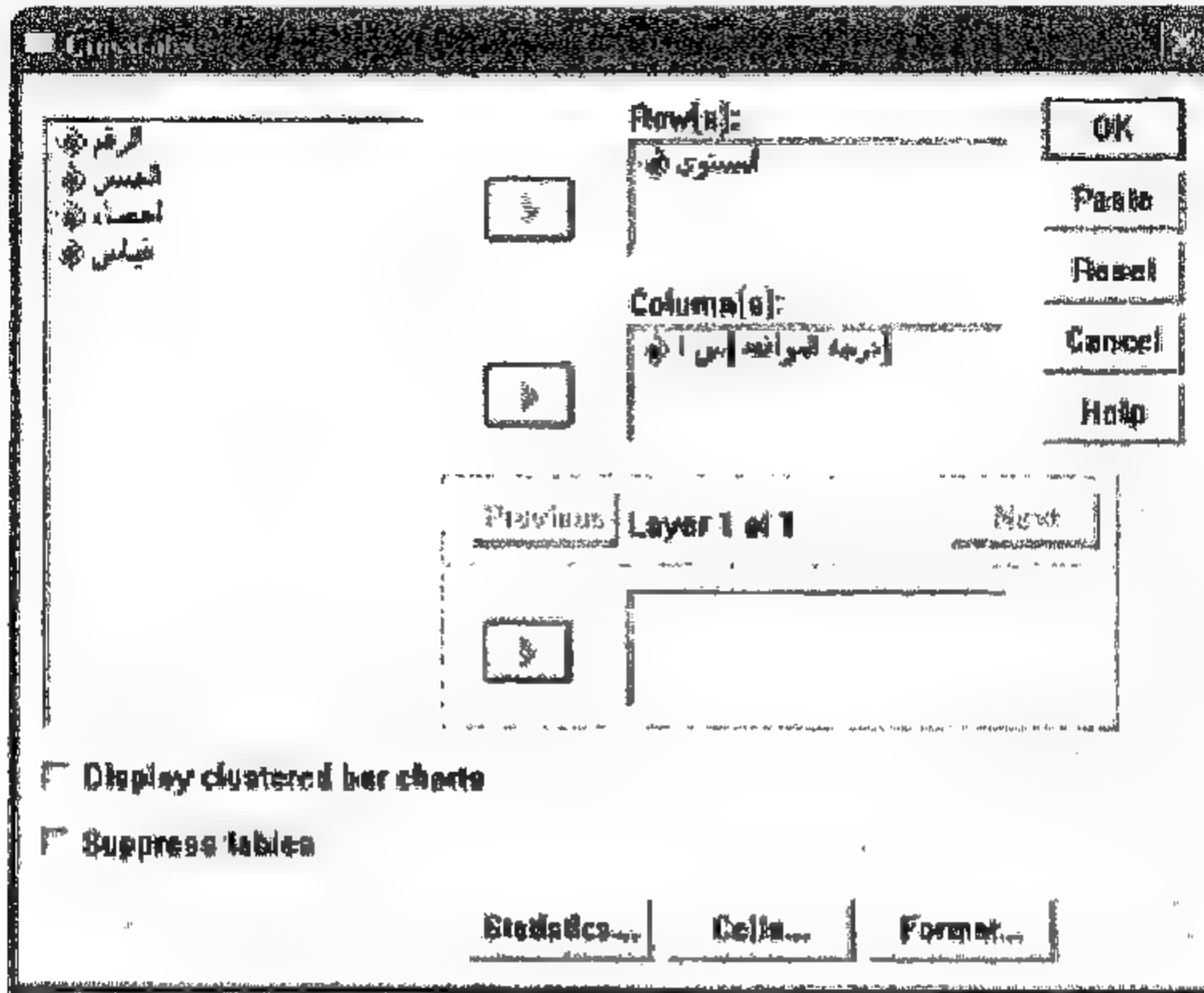
		الاحصاء	القياس
الاحصاء	Pearson Correlation	1	.257
	Sig. (2-tailed)		.473
	N	10	10
القياس	Pearson Correlation	.257	1
	Sig. (2-tailed)	.473	.
	N	10	10

معامل الارتباط بيرسون بين الاحصاء والقياس = 0.257

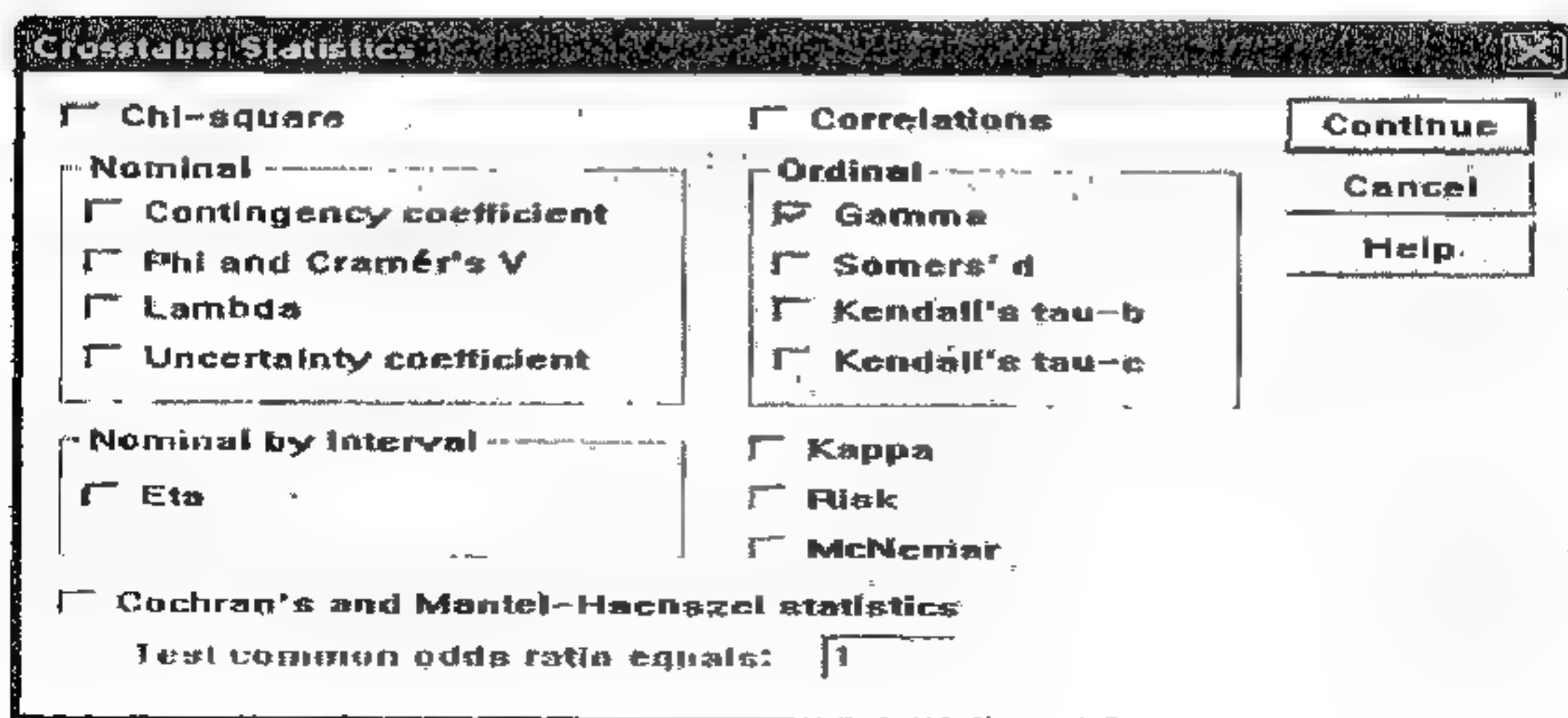
وهو غير دال احصائياً لأن $\text{Sig}=0.473$ وهي اكبر من مستوى الدلالة $(\alpha=0.05)$

9- هل هناك علاقة بين مستوى الطالب ودرجة موافقته وهل هو دال على مستوى $(\alpha=0.05)$ ولصالح من؟

لعمل معامل ارتباط Analyze – Descriptive Statistics – Crosstabs...



نضغط زر Statistics...



Crosstabs

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
المستوى "درجة الموافقة"	10	100.0%	0	.0%	10	100.0%

Crosstabulation

Count		درجة الموافقة					Total
		مستوى عدم الموافقة	مستوى متروك	متوسط	مستوى موافقة	مستوى موافقة عالية	
		1			1		2
مستوى موافقة	مستوى موافقة	1	1		1	1	4
مستوى موافقة	مستوى موافقة			2	1		3
مستوى موافقة	مستوى موافقة				1		1
Total		2	1	2	3	2	10

Symmetric Measures

	Value	Asymp. Std. Error ^a	Approx. T ^b	Approx. Sig.
Ordinal by Ordinal Gamma	.400	.303	1.222	.222
N of Valid Cases	10			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Correlations

			المستوى	درجة الموافقة
Spearman's rho	المستوى	Correlation Coefficient	1.000	.361
		Sig. (2-tailed)	.	.306
		N	10	10
	درجة الموافقة	Correlation Coefficient	.361	1.000
		Sig. (2-tailed)	.306	
		N	10	10

معامل الارتباط جاما بين مستوى الطالب ودرجة الموافقة = 0.400

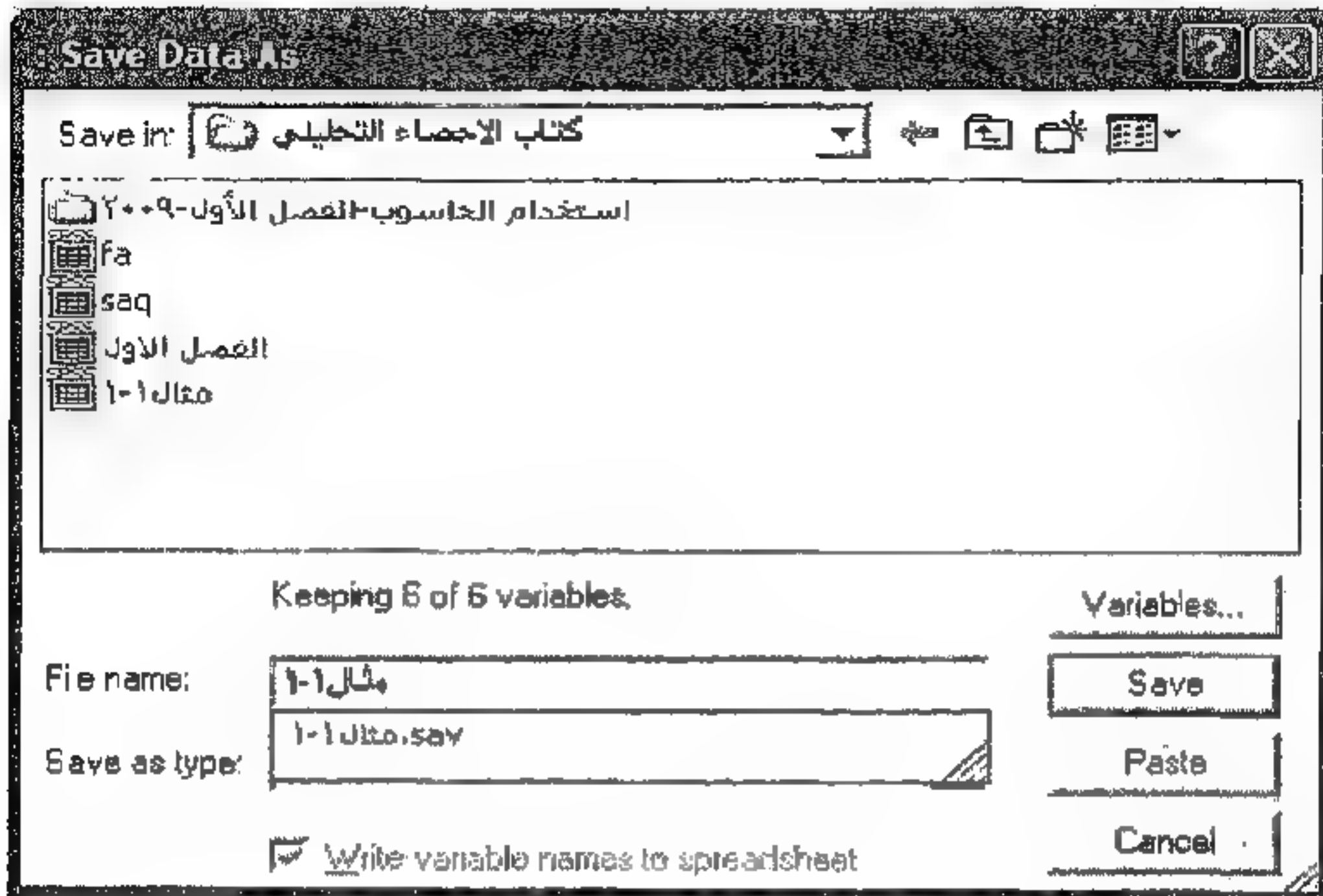
وهو غير دال احصائياً لأن Sig=0.222 وهي اكبر من مستوى الدلالة ($\alpha=0.05$)

معامل الارتباط سبيرمان بين مستوى الطالب ودرجة الموافقة = 0.361

وهو غير دال احصائياً لأن Sig=0.306 وهي اكبر من مستوى الدلالة ($\alpha=0.05$)

10- اعمل على حفظ الملف تحت اسم مثال 1-1

لحفظ الملف من قائمة File – Save As....



ثم نضغط زر Save

مثال 1-2:

في اختبار لمادة الحاسوب الذي يتكون من (15) سؤال من نوع الاختيار من متعدد والذي اجاب عليه (20) طالب، والمطلوب منك تحليل هذا الاختبار والتعليق على نتائجه؟

رقم الطالب: التخصص: 1-علمي 2-أدبي 3-معلوماتية 4-مهني

الجنس: 1-ذكر 2-أنثى

المادة: الحاسوب. المستوى الثالث: تراسل البيانات والشبكات

	q1	q2	q3	q4	q5	q6	q7	q8	q9	q10	q11	q12	q13	q14	q15
1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1
2	2	1	0	1	1	1	1	1	1	1	1	1	1	1	1
3	3	1	0	0	1	1	1	1	1	1	1	1	1	1	1
4	4	1	0	0	1	1	1	1	1	1	1	1	1	1	1
5	5	1	0	0	1	1	1	1	1	1	1	0	1	1	1
6	6	1	0	1	1	0	1	1	1	1	1	0	1	1	1
7	7	1	0	0	1	0	1	1	1	0	1	0	1	1	1
8	8	1	0	0	1	0	1	1	1	0	1	1	1	1	1
9	9	1	0	1	1	0	1	1	1	0	1	1	1	1	1
10	10	1	0	0	1	0	1	1	1	1	1	1	1	1	1
11	11	2	0	1	1	1	1	1	1	1	1	1	1	1	1
12	12	2	0	1	1	1	1	1	1	1	1	1	1	1	1
13	13	2	0	1	1	0	1	1	1	1	1	1	1	1	1
14	14	2	0	0	0	0	0	1	1	0	0	0	1	1	1
15	15	2	0	0	0	0	0	1	1	1	1	1	1	1	1
16	16	2	1	0	0	0	0	1	1	1	1	1	1	1	1
17	17	3	1	0	1	1	0	0	1	1	1	1	1	1	1
18	18	2	0	0	0	1	0	1	1	1	1	0	0	1	0
19	19	2	1	1	1	1	1	1	1	1	0	1	1	1	1
20	20	2	1	0	0	1	0	1	1	1	0	0	0	1	1

* أوجد ما يلي: مع إعطاء التعليق المناسب.

- 1- اعمل على ترميز المتغيرات الواردة في السؤال Coding
- 2- اعمل على ادخال البيانات الخاصة بالاختبار كما هو مبين ادناه
- 3- علامة كل طالب على الاختبار.
- 4- الوسط الحسابي والانحراف المعياري لكل سؤال في الاختبار.
- 5- الوسط الحسابي والانحراف المعياري للاختبار.
- 6- معامل الثبات للاختبار.
- 7- معامل الارتباط بين السؤال الاول والمجموع وهل هو دال على مستوى ($\alpha=0.01$)
- 8- اعمل على حفظ الملف تحت اسم مثال 1-2

1- اعمل على ترميز المتغيرات الواردة في السؤال Coding

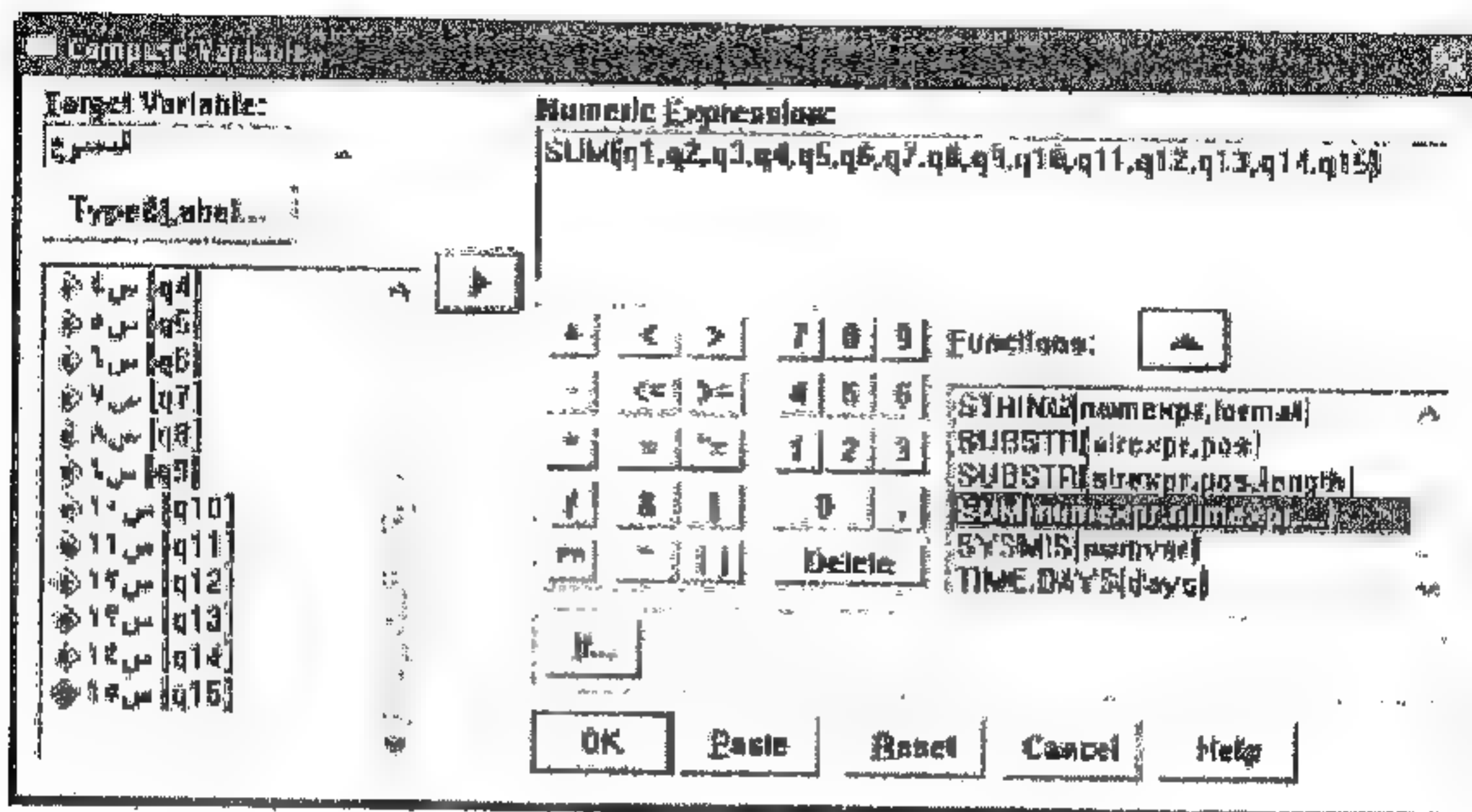
	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	slno	Numeric	3	0	رقم الطالب	None	None	4	Center	Scale
2	sex	Numeric	1	0	الجنس	1. ذكر، 2. أنثى	None	3	Center	Nominal
3	spec	Numeric	1	0	المتخصص	1. أعطى، 2. لم يعط	None	3	Center	Nominal
4	q1	Numeric	1	0	س1	None	None	3	Center	Scale
5	q2	Numeric	1	0	س2	None	None	3	Center	Scale
6	q3	Numeric	1	0	س3	None	None	4	Center	Scale
7	q4	Numeric	1	0	س4	None	None	4	Center	Scale
8	q5	Numeric	1	0	س5	None	None	4	Center	Scale
9	q6	Numeric	1	0	س6	None	None	4	Center	Scale
10	q7	Numeric	1	0	س7	None	None	4	Center	Scale
11	q8	Numeric	1	0	س8	None	None	4	Center	Scale
12	q9	Numeric	1	0	س9	None	None	4	Center	Scale
13	q10	Numeric	1	0	س10	None	None	4	Center	Scale
14	q11	Numeric	1	0	س11	None	None	3	Center	Scale
15	q12	Numeric	1	0	س12	None	None	4	Center	Scale
16	q13	Numeric	1	0	س13	None	None	3	Center	Scale
17	q14	Numeric	1	0	س14	None	None	4	Center	Scale
18	q15	Numeric	1	0	س15	None	None	4	Center	Scale

2- اعمل على ادخال البيانات الخاصة بالاختبار كما هو مبين ادناه

	slno	sex	spec	q1	q2	q3	q4	q5	q6	q7	q8	q9	q10	q11	q12	q13	q14	q15
1	1	1	3	1	1	0	0	1	1	1	1	0	1	1	1	1	1	1
2	2	1	3	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1
3	3	1	3	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1
4	4	1	3	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1
5	5	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1
6	6	1	1	1	1	0	1	1	0	1	1	1	1	1	0	1	1	1
7	7	1	1	1	1	0	0	1	0	1	1	1	1	1	0	1	1	1
8	8	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	1
9	9	1	1	1	1	0	1	1	0	1	1	1	1	0	1	1	1	1
10	10	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	1
11	11	2	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
12	12	2	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
13	13	2	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1
14	14	2	1	0	1	0	0	0	0	0	1	1	0	0	1	1	1	1
15	15	2	1	0	0	1	0	0	0	1	1	1	1	1	1	1	1	1
16	16	2	2	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1
17	17	2	2	1	1	0	1	1	0	0	1	1	1	1	1	1	1	1
18	18	2	2	0	0	0	1	0	1	1	1	1	1	1	0	0	1	0
19	19	2	2	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1
20	20	2	2	1	1	0	0	1	0	1	1	1	1	0	0	0	1	1

3- علامة كل طالب على الاختبار.

Transform – Compute...

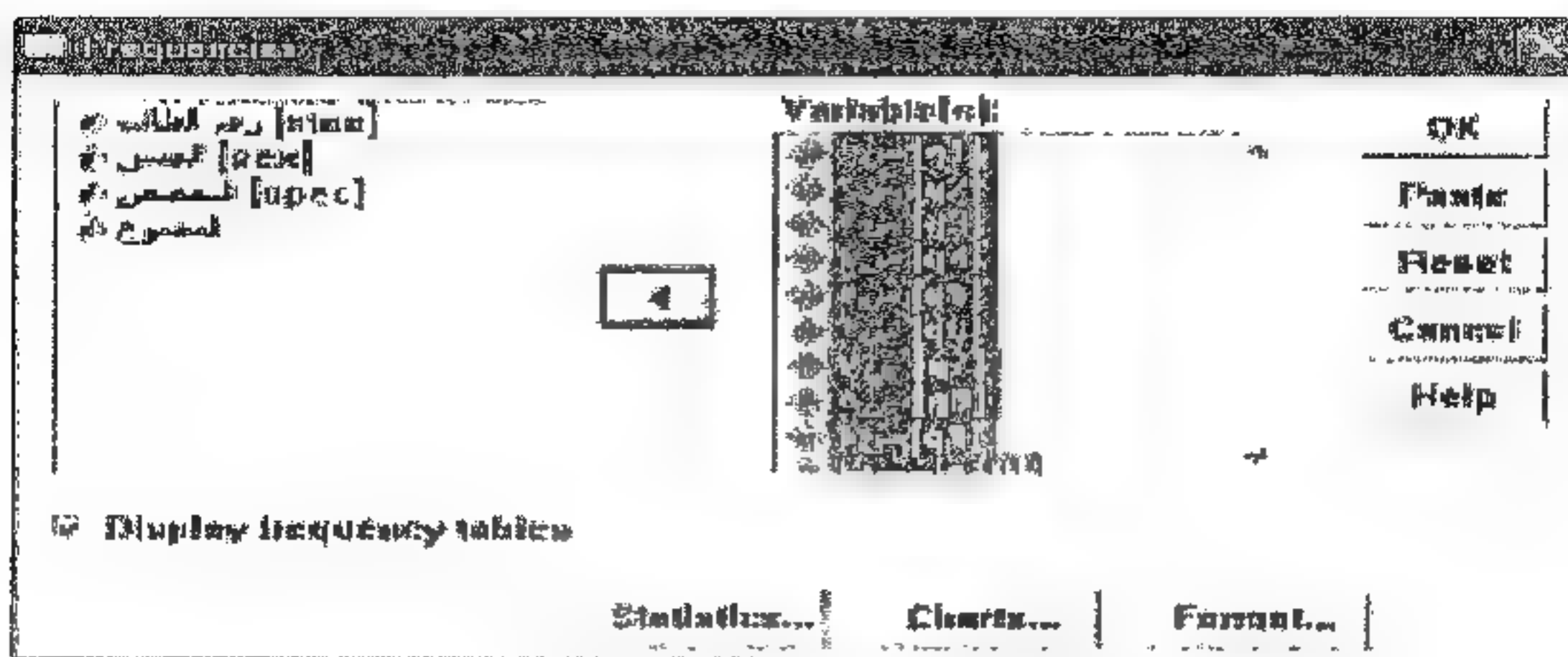


ثم نضغط Ok

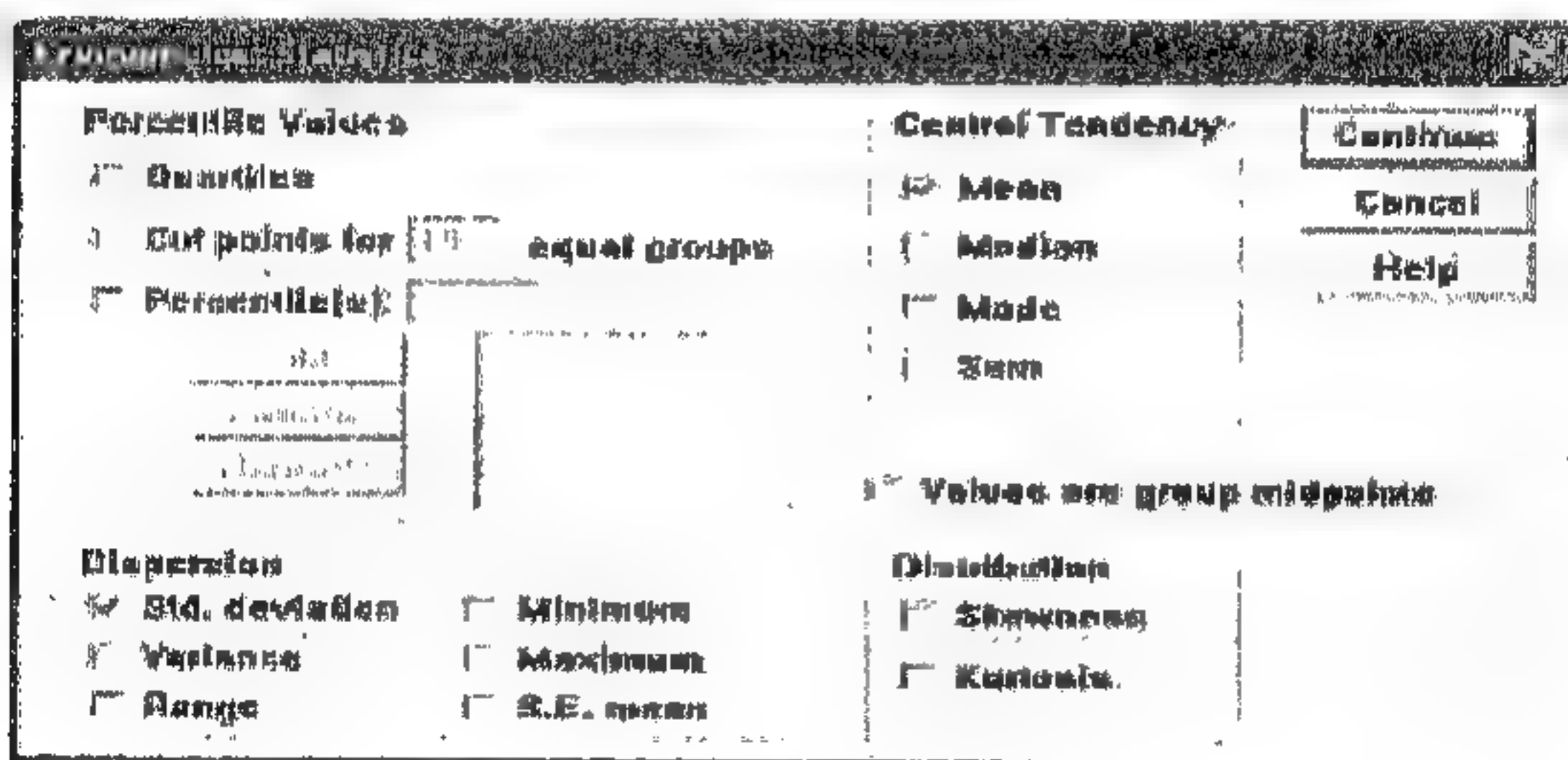
	الدرجة	q1	q2	q3	q4	q5	q6	q7	q8	q9	q10	q11	q12	q13	q14	q15
1	13.00	1	1	3	1	1	0	0	1	1	1	1	1	1	1	1
2	14.00	2	1	3	1	1	0	1	1	1	1	1	1	1	1	1
3	13.00	3	1	3	1	1	0	0	1	1	1	1	1	1	1	1
4	13.00	4	1	3	1	1	0	0	1	1	1	1	1	1	1	1
5	13.00	5	1	1	1	1	0	0	1	1	1	1	0	1	1	1
6	12.00	6	1	1	1	1	0	1	1	0	1	1	0	1	1	1
7	10.00	7	1	1	1	1	0	0	1	0	1	1	0	1	1	1
8	10.00	8	1	1	1	0	0	0	1	0	1	1	1	1	1	1
9	11.00	9	1	1	1	0	0	1	1	0	1	1	1	1	1	1
10	12.00	10	1	1	1	1	0	0	1	0	1	1	1	1	1	1
11	14.00	11	2	1	1	1	0	1	1	1	1	1	1	1	1	1
12	14.00	12	2	1	1	1	0	1	1	1	1	1	1	1	1	1
13	13.00	13	2	1	1	1	0	1	1	0	1	1	1	1	1	1
14	6.00	14	1	1	0	1	0	0	0	0	1	1	0	0	1	1
15	10.00	15	2	1	0	0	1	0	0	0	1	1	1	1	1	1
16	12.00	16	2	2	1	1	0	0	0	1	1	1	1	1	1	1
17	12.00	17	2	2	1	1	0	1	1	0	0	1	1	1	1	1
18	8.00	18	2	2	0	0	0	1	0	1	1	1	0	0	1	0
19	14.00	19	2	2	1	1	1	1	1	1	1	0	1	1	1	1
20	6.00	20	2	2	1	0	0	1	0	1	1	1	0	0	1	1

4- الوسط الحسابي والانحراف المعياري لكل سؤال في الاختبار.

Analyze – Descriptive Statistics - Frequencies... لعمل إحصاءات وصفية...



نضبط زر Statistics...

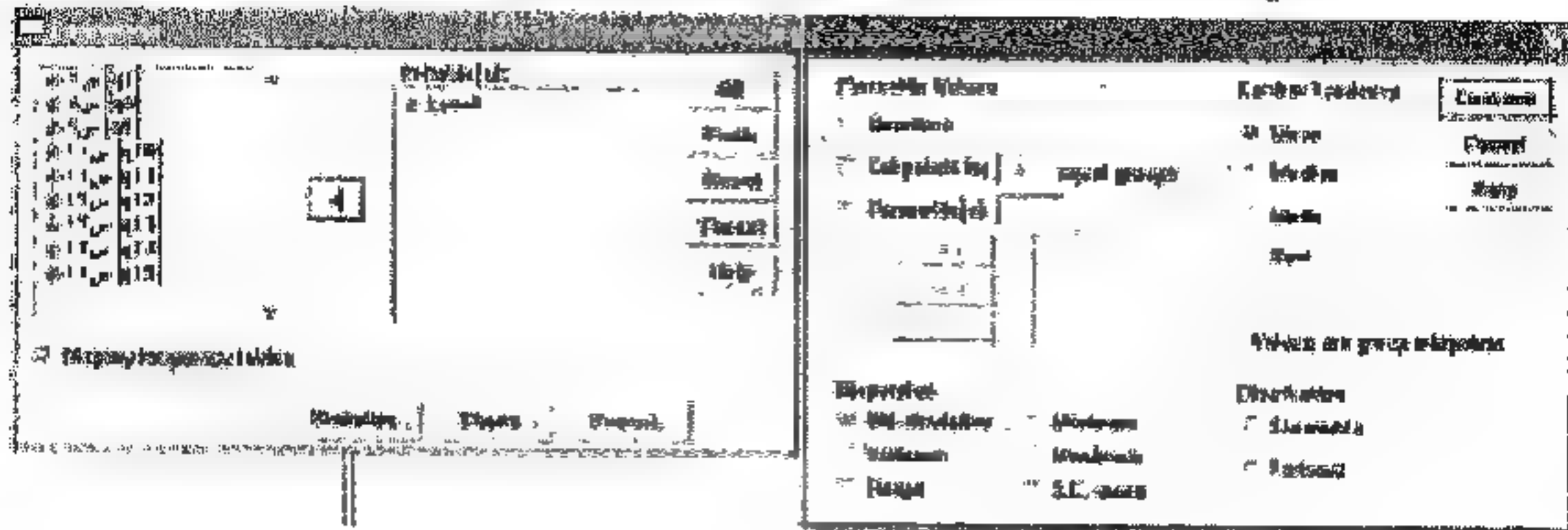


ضغط زر Continue... ، ثم اضغط زر Ok

Table 1.1.1.1												
Year	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971
1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972
1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972
1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972

		1966	1967	1968	1969	1970
FI	74.41	72	72.1	70	70	70
	W. 10.00	11	11	12	11	11
Mean's	84.41	83	83.1	81	80.5	80
Std. Dev.	5.11	5.03	5.04	5.00	5.00	5.00

5- الوسط الحسابي والانحراف المعياري للاختبار.



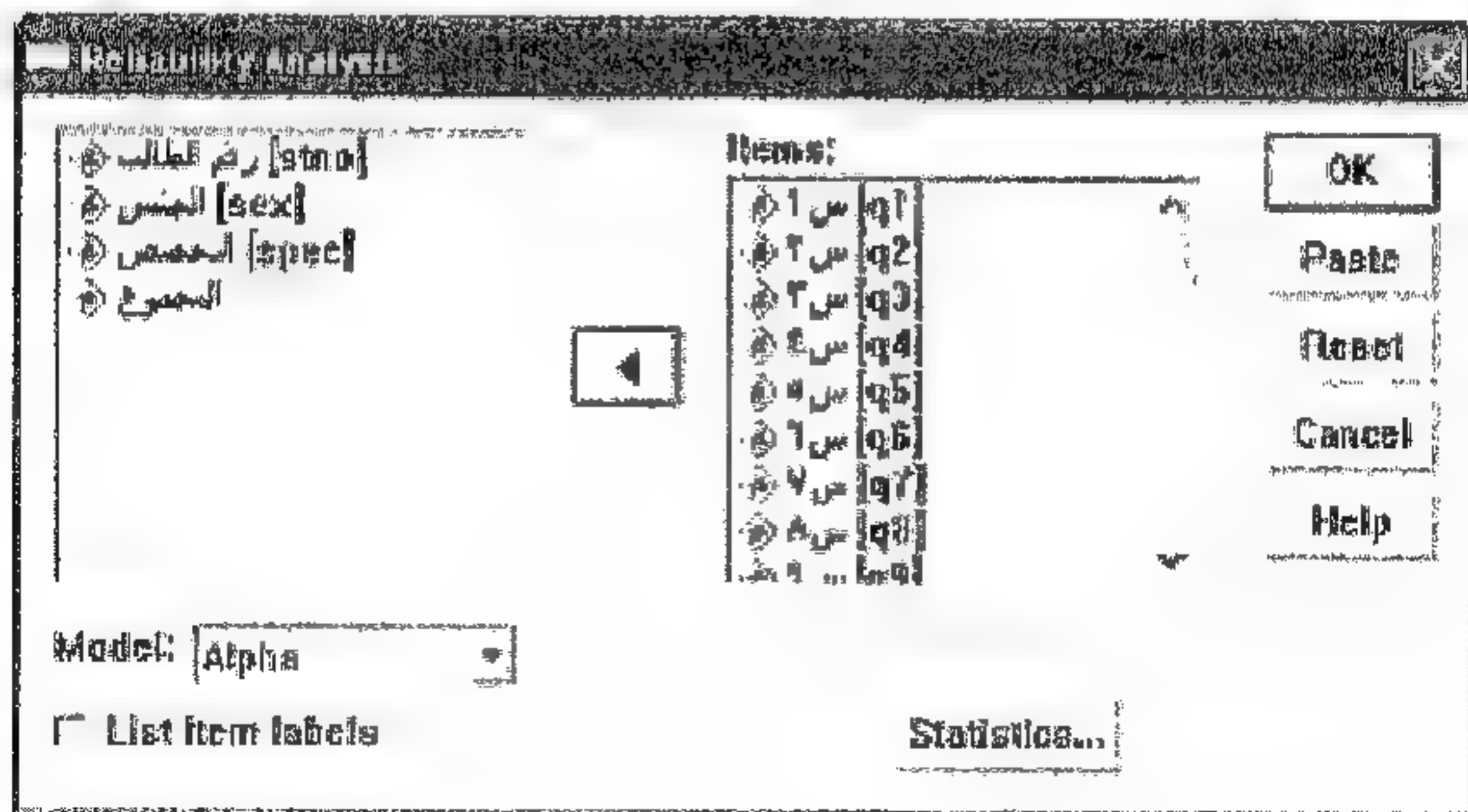
نضغط زر Continue... ، ثم نضغط زر Ok

Statistics

المجموع		
N	Valid	20
	Missing	0
Mean		11.6000
Std. Deviation		2.18608

6- معامل الثبات للاختبار.

Analyze – Scale – Reliability Analysis



ثم نضغط زر Ok

الفصل الأول: مفاهيم أساسية في الإحصاء

→ Reliability

***** Method 1 (space saved) will be used for this analysis *****

RELIAbility ANALYSIS - SPSS (ALPHA)				
	Item	Mean	Std. Dev.	Alpha
1.	Q1	.6500	.3563	.60,0
2.	Q2	.6000	.4104	.60,0
3.	Q3	.3500	.3483	.60,0
4.	Q4	.4500	.3104	.60,0
5.	Q5	.6000	.4104	.60,0
6.	Q6	.4500	.3104	.60,0
7.	Q7	.3500	.3483	.60,0
8.	Q8	1.0000	.0000	.60,0
9.	Q9	1.0000	.0000	.60,0
10.	Q10	.6000	.3483	.60,0
11.	Q11	.4500	.3104	.60,0
12.	Q12	.7000	.4702	.60,0
13.	Q13	.4500	.3483	.60,0
14.	Q14	1.0000	.0000	.60,0
15.	Q15	.9500	.2236	.60,0

Split Reliability Coefficients

N of Cases = 20,0

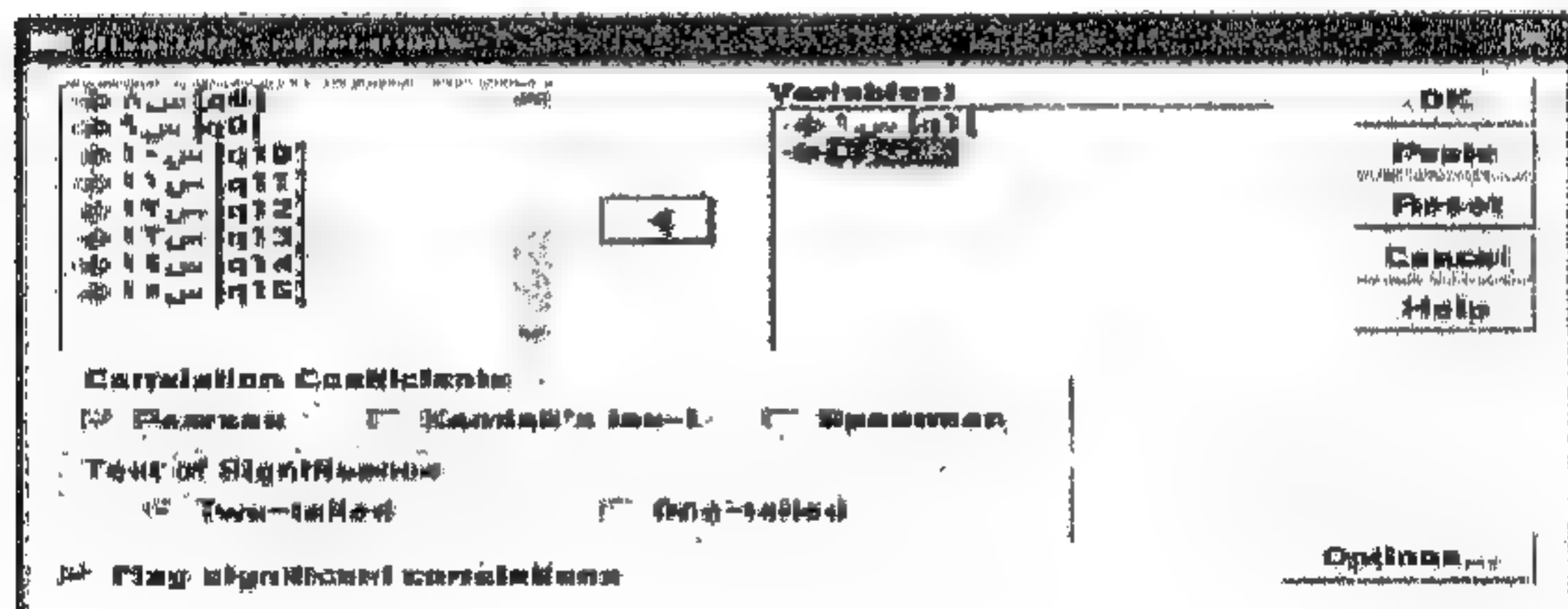
N of Items = 15

Alpha = .6078

7- معامل الارتباط بين السؤال الاول والجموع وهل هو دال على مستوى ($\alpha=0.01$).

Analyze - Correlate - Bivariate...

لحساب معامل الارتباط



ثم نضغط زر Ok

Correlations

		س1	الجموع
س1	Pearson Correlation	1	.710**
	Sig. (2-tailed)	.	.000
	N	20	20
الجموع	Pearson Correlation	.710**	1
	Sig. (2-tailed)	.000	
	N	20	20

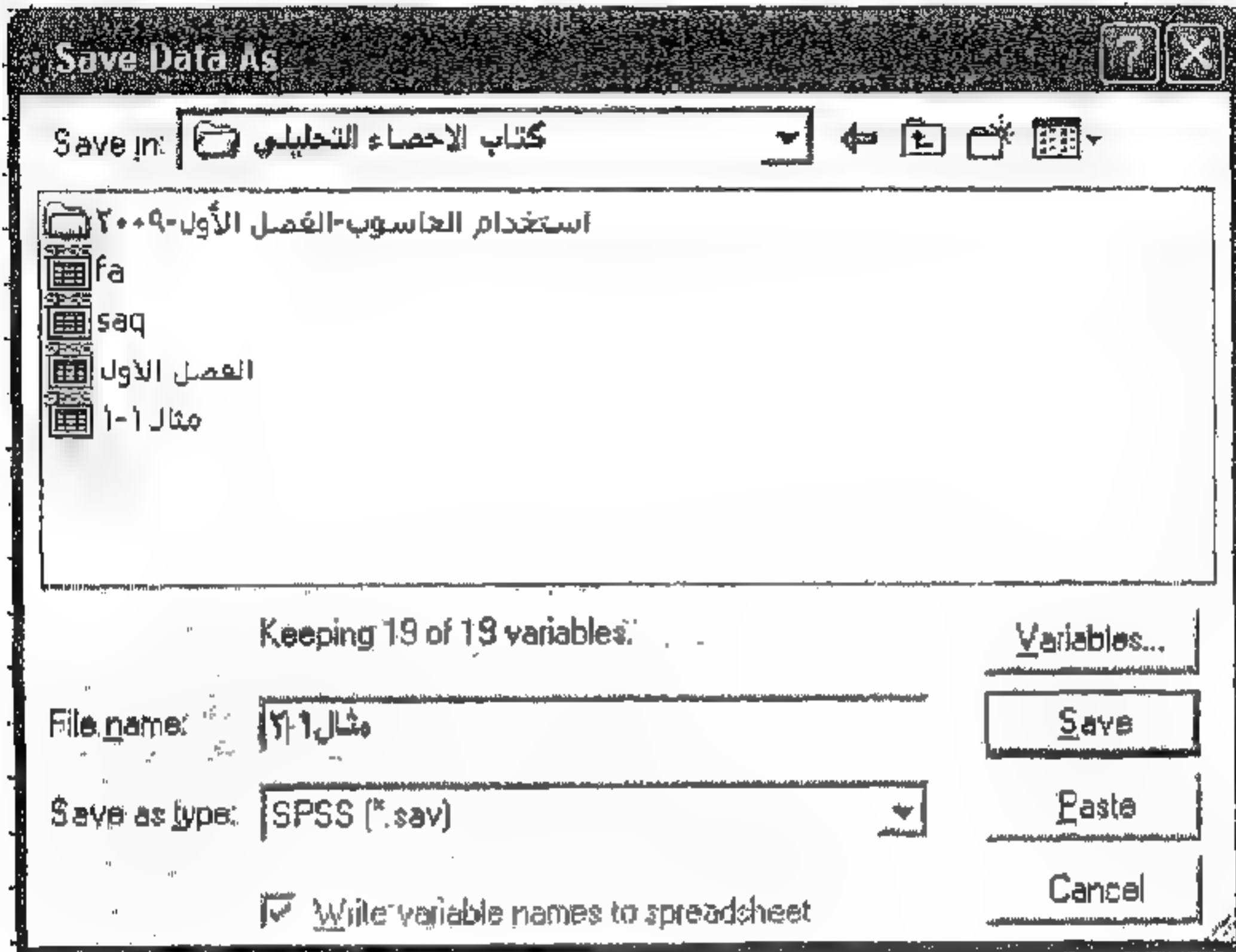
** . Correlation is significant at the 0.01 level

معامل الارتباط بيرسون بين السؤال الاول والجموع = .7100

وهو ذو دلالة احصائية لأن $Sig=0.000$ وهي أقل من مستوى الدلالة ($\alpha=0.01$)

8- اعمل على حفظ الملف تحت اسم مثال 1-2

لحفظ الملف من قائمة File – Save As....



ثم نضغط زر Save

6-1 تمارين Exercise

س1: ضع رمز الإجابة الصحيحة في المربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										
الرقم	11	12	13	14	15	16	17	18	19	20
الإجابة										
الرقم	21	22	23	24	25	26	27	28	29	30
الإجابة										

1- البيانات المبوبة هي البيانات:

- أ- الأولية ب- المجدولة ج- الرقمية د- غير الرقمية

2- البيانات النوعية هي التي تصف الظاهرة بشكل:

- أ- أولي ب- نهائي ج- غير رقمي د- رقمي

3- إن علامة الطالب هي مثال على البيانات:

- أ- المبوبة ب- رمزي ج- النوعية د- الكمية

4- إن جنس الطالب هو مثال على البيانات:

- أ- المنفصلة ب- المتصلة ج- النوعية د- الكمية

5- مؤشر المجتمع يسمى:

- أ- إحصائي ب- معلم ج- مؤشر مجتمع د- مؤشر

6- مؤشر العينة يسمى:

- أ- مؤشر ب- معلم عينة ج- معلم د- إحصائي

7- إن الرمز التالي S_x معناه:

- أ- تباين مجتمع ب- انحراف معياري مجتمع
ج- تباين عينة د- انحراف معياري عينة

- 8- تصنف المتغيرات حسب طبيعة المعلومات التي يؤديها القياس إلى:
- أ- اسمية، ترتيبية، فئوية، نسبية.
 ب- نوعية، كمية.
 ج- مستقلة، تابعة.
 د- رقمية، رمزية.
- 9- تسمى المتغيرات التي لها عدد فئات محدد من دون أي وزن لهذه الفئات ولا يوجد أفضلية لأحدها على الآخر بالمتغيرات.
- أ- اسمية
 ب- ترتيبية
 ج- فئوية
 د- نسبية.
- 10- إن متغير الجنس الذي يصنف المجتمع إلى فئتين هما الذكور والإناث هو مثال على المتغيرات.
- أ- اسمية
 ب- ترتيبية
 ج- فئوية
 د- نسبية.
- 11- الأرقام التي ليس لها معنى حقيقي ولا يمكن إجراء العمليات الحسابية عليها هي مثال على المتغيرات.
- أ- اسمية
 ب- ترتيبية
 ج- فئوية
 د- نسبية.
- 12- تسمى المتغيرات التي لها عدد محدد من الفئات يمكن ترتيبها تصاعدياً أو تنازلياً، ولا يمكن تحديد الفروق بدقة بين القيم المختلفة بالمتغيرات.
- أ- اسمية
 ب- ترتيبية
 ج- فئوية
 د- نسبية.
- 13- إذا كانت A أكبر من B ولكن لا نستطيع معرفة كم يكبر A عن B هي مثال على المتغيرات.
- أ- اسمية
 ب- ترتيبية
 ج- فئوية
 د- نسبية.
- 14- المتغيرات الكمية التي يمكن إجراء العمليات الحسابية عليها وذلك دون أن تتأثر المسافة النسبية بين قيمها، ويميز هذا المتغير من خلال قيمة الصفر التي لا تعني عدم توافر تلك الصفة تسمى بالمتغيرات.
- أ- اسمية
 ب- ترتيبية
 ج- فئوية
 د- نسبية.

15- إذا كانت علامة جمال في مادة اللغة العربية أكثر من علامة أيمن، وأن علامة أيمن أكثر من علامة محمد، فإننا نعرف هذه المتغيرات بالمتغيرات.

أ- نسبية ب- فتوية ج- ترتيبية د- اسمية

16- المتغيرات الكمية التي ليس لها فئات محددة ولكن الصفر فيها يمثل عدم توفر الصفة، مثل المتغيرات الزمنية تسمى بالمتغيرات.

أ- اسمية ب- ترتيبية ج- فتوية د- نسبية.

17- إذا أراد مدرس أن يبحث عن أثر عدد ساعات الدراسة على تحصيل الطالب في مبحث معين، فإن المتغير المستقل هو؟

أ- المدرس ب- ساعات الدراسة ج- التحصيل د- الطالب.

18- إذا أراد مدرس أن يبحث عن أثر عدد ساعات الدراسة على تحصيل الطالب في مبحث معين، فإن المتغير التابع هو؟

أ- المدرس ب- ساعات الدراسة ج- التحصيل د- الطالب.

19- العينة التي يتم اختيارها بعناية وبصورة غير عشوائية تسمى بالعينة.

أ- القصدية ب- الطبقية ج- العنقودية د- المنتظمة.

20- العينة التي يتم اختيار أفرادها من المجتمع بحيث يكون لأي فرد من الأفراد الفرصة نفسها للظهور في هذه العينة تسمى بالعينة.

أ- القصدية ب- العشوائية البسيطة ج- العنقودية د- المنتظمة.

21- إذا كان حجم المجتمع كبير جداً، وقسم المجتمع إلى مجموعات صغيرة، ثم اختبرت عينة عشوائية من هذه المجموعات الصغيرة تسمى العينة بالعينة.

أ- القصدية ب- العشوائية البسيطة ج- العنقودية د- المنتظمة.

22- العينة التي تمثل المجتمع الإحصائي تمثيلاً صادقاً وتتفق مقاييسها الإحصائية مع مقاييس المجتمع ويتم اختيارها بصورة تابعة تسمى بالعينة.

أ- القصدية ب- العشوائية البسيطة ج- العنقودية د- المعيارية.

23- إذا كان السبب في أخطاء العينات هو خطأ في طريقة اختيار العينة، تسمى أخطاء العينة في هذه الحالة بأخطاء:

أ- عشوائية ب- التحيز ج- العينة د- المعطيات.

24- إذا كان السبب في أخطاء العينات هو زيادة أو نقص في البيانات، تسمى أخطاء العينة في هذه الحالة بأخطاء:

أ- عشوائية ب- التحيز ج- العينة د- المعطيات.

25- نتائج العينات تتعرض لمجموعة من الأخطاء منها ما يلي:

أ- المنتظمة ب- التحيز ج- العينة د- جميع ما ذكر

26- في تجربة لمعرفة أثر تخصص الطالب على تحصيله في مساق مهارات الحاسوب، ينظر إلى متغير تخصص الطالب كمتغير:

أ- مستقل ب- تابع ج- نسبي د- فئوي

27- رقم الشعبة في مساق متعدد الشعب مثال على المتغير:

أ- رتي ب- الاسمي ج- نسبي د- فئوي

28- المتغير الذي لمقياسه صفر مطلق هو المتغير:

أ- رتي ب- الاسمي ج- نسبي د- فئوي

29- التعريف التالي: "جزء من المجتمع تتم دراسة الظاهرة من خلال المعلومات عنها".

أ- الإحصاء ب- المجتمع ج- العينة د- القياس

30- يختلف الإحصاء الوصفي عن الإحصاء التحليلي بشكل أساسي في:

أ- الطريقة ب- الدقة ج- التعميم د- نوع البيانات

س2: ضع رمز الإجابة الصحيحة في المربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										
الرقم	11	12	13	14	15	16	17	18	19	20
الإجابة										
الرقم	21	22	23	24	25	26				
الإجابة										

1- الوسط الحسابي للقيم 0، 5، 1، -2، 6 هو:-

أ) 0 ب) -2 ج) 2 د) 5

2- ما قيمة الوسيط لمجموعة القيم التالية: 16، 18، 20، 24، 32، 22

أ) 11 ب) 16 ج) 20 د) 21

3- إذا كان وسيط (15) علامة هو (60) وعدلت العلامات حسب المعادلة ص = -0.9س + 3

حيث س العلامة السابقة، ص العلامة الجديدة، فإن الوسيط الجديد

أ) 54 ب) 15 ج) 57 د) 17.5

4- يعرف المئين 80 بأنه:

أ) قيمة في التوزيع = 80

ب) قيمة في التوزيع أكبر من 80% من مجموع القيم

ج) قيمة في التوزيع تزيد عن 80% من القيم

د) قيمة في التوزيع فوق الوسيط بمقدار 30 علامة

5- أي من المقاييس التالية يتأثر بالقيم المتطرفة :

أ) الوسط ب) الوسيط ج) المنوال د) لا شيء مما ذكر

6- أي من المقاييس التالية يتأثر بالتحويلات الخطية (الجمع والطرح والضرب والقسمة):

أ) الوسط ب) الوسيط ج) المنوال د) جميع ما ذكر

7- أي من المقاييس التالية لا يعتمد على قيم البيانات إنما يعتمد على ترتيبها وموقعها:

أ) الوسط ب) الوسيط ج) المنوال د) لا شيء مما ذكر

استعن بالجدول للإجابة عن الأسئلة التالية: - (8، 9، 10، 11، 12)

التكرار التراكمي	الحدود الفعلية	المراكز* التكرار	المراكز	تكرار	الفئات
1	-0.5 – 4.5	2	2	1	00-04
3	4.5 – 9.0	14	7	2	05-09
6	9.5 – 14.5	36	12	3	10-14
8	14.5 – 19.5	34	17	2	15-19
10	19.5 – 24.5	44	22	2	20-24
		130		10	المجموع

8- مركز الفئة الرابعة هو:-

أ) 12 ب) 17 ج) 22 د) لا يمكن معرفته

9- الوسط الحسابي للتوزيع (\bar{X}) -

أ) 6 ب) 10 ج) 13 د) 60

10- المئين 60 للتوزيع -

أ) 14.5 ب) 12.5 ج) 9.5 د) 3

11- الحدود الفعلية للفئة الخامسة هي :

أ) 24-20 ب) 24.5-19.5

ج) 23.5-20.5 د) 24.5-20.5

12- طول الفئة للفئة الرابعة يساوي :

أ) 5 ب) 4 ج) 3 د) لا شيء مما ذكر

13- ما المنوال للقيم التالية: 4، 7، 2، 9، 2، 11

أ) 2 ب) 7 ج) 9 د) 11

- 14- إن مقاييس التزعة المركزية تتأثر بالعمليات التالية:
- أ) الجمع ب) الطرح ج) الضرب د) جميع ما ذكر
- 15- منوال التوزيع التالي 3، 2، 2، 6، 7، 7 هو
- أ) 2 ب) 7 ج) 3 د) 2، 7
- 16- أكثر مقاييس التزعة المركزية تأثراً بالقيم المتطرفة هو
- أ) الوسط ب) الوسيط ج) المنوال د) المئين
- 17- أفضل مقاييس التزعة المركزية لوصف المتغيرات الاسمية هو
- أ) الوسط ب) الوسيط ج) المنوال د) المئين
- 18- أقل مقاييس التزعة المركزية تأثراً بتقلبات العينة هو
- أ) الوسط ب) الوسيط ج) المنوال د) المئين
- 19- أي مقاييس التزعة المركزية يقسم التوزيع إلى قسمين متساويين مهما كان شكله
- أ) الوسط ب) الوسيط ج) المنوال د) المئين
- 20- إذا علمت أن الوسط الحسابي لعلامات طلاب صف ما هي 10 والانحراف المعياري 2 فإذا أضيفت 3 علامات لكل طالب فإن الوسط الحسابي للعلامات بعد الإضافة هو
- أ) 10 ب) 2 ج) 3 د) 13
- 21- إن الوسط الحسابي لخمس قيم ثلاث منها متوسطها 4 والباقي متوسطها 5 هو:
- أ) 10 ب) 12 ج) 24 د) 4.4
- 22- احسب الوسط الحسابي للقيم التالية : 12، 15، 14، 15، 22
- أ) 78 ب) 15.6 ج) 15 د) 3
- 23- احسب الوسيط للقيم التالية : 12، 15، 14، 15، 22
- أ) 78 ب) 15.6 ج) 15 د) 3

24- احسب المتوسط الحسابي للقيم التالية : 12، 15، 14، 15، 22

- أ) 78 ب) 15.6 ج) 15 د) 3

25- إذا علمت أن الوسط الحسابي لعلامات طلاب مساق الإحصاء هو 20 بينما كان

الوسط الحسابي لعلامات الطالبات في نفس الصف هي 15 فإذا كان عدد الطلاب في الشعبة 15 وعدد الطالبات 10 فإن الوسط الحسابي لعلامات الشعبة كاملة هي

- أ) 18 ب) 20 ج) 15 د) 17.5

26- إذا علمت أن الوسط الحسابي لعلامات 30 طالباً في امتحان مادة الإحصاء هو 70 ، ثم

انسحب طالب علامته 80 ، فإن الوسط الحسابي للعلامات بعد الانسحاب هو

- أ) 70 ب) 69.56 ج) 69.65 د) 80

س3: ضع رمز الإجابة الصحيحة في المربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										

1- ما قيمة المدى للعلامات التالية 4، 12، -1، 8، 11:-

- أ) -1 ب) 11 ج) 13 د) 7

2- الانحراف المتوسط للقيم 1، 2، 3 هو:-

- أ) $\sqrt{2}$ ب) $\sqrt{3}$ ج) $\sqrt{\frac{2}{3}}$ د) $\frac{3}{2}$

3- الانحراف المعياري للقيم 5، 4، 3، 2، 1 هو:-

- أ) $\sqrt{2}$ ب) $\sqrt{3}$ ج) $\sqrt{5}$ د) $\sqrt{10}$

4- إذا كان الانحراف المعياري لـ (10) قيم هو 9 فإن التباين =

- أ) 3 ب) $\sqrt{10}$ ج) 81 د) 100

5- إن مقاييس التشتت تتأثر بالعمليات التالية :

- أ) الجمع ب) الطرح ج) الضرب د) لاشيء مما ذكر

* إذا علمت أن درجات 10 طلاب في مادة الإحصاء هي كما يلي:

50 ، 70 ، 85 ، 80 ، 40 ، 65 ، 60 ، 90 ، 85 ، 75 أجب عن الأسئلة من

10-6

6- أ أن المدى لدرجات الطلاب هو:

أ) 50 ب) 25 ج) 26 د) 40

7- أ أن الانحراف المتوسط لدرجات الطلاب هو:

أ) 130 ب) 13 ج) 10 د) 50

8- أ أن الانحراف المعياري لدرجات الطلاب هو:

أ) 900 ب) 10 ج) 90 د) $\sqrt{90}$

9- أ أن الثباين لدرجات الطلاب هو:

أ) 900 ب) 10 ج) 90 د) $\sqrt{90}$

10- أ أن معامل الاختلاف لدرجات الطلاب هو:

أ) $\sqrt{70}$ ب) $\sqrt{90}/70$ ج) 90 د) $\sqrt{90}$

س4: ضع رمز الإجابة الصحيحة في الربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										
الرقم	11	12	13	14						
الإجابة										

1- حسب معامل الارتباط بين Y , X فكان -0.65 ، عرف المتغيران ، $L = 2X - 0.2$

$M = -3X + 0.3$ فإن معامل الارتباط الجديد بين L , M هو:

أ- (-0.35) ب- (-0.85) ج- (-0.65) د- 0.65

2- إذا كان $\sum(x-x') = 27$ ، وكان $\sum(x-x')^2 = 36$ ، وكان $\sum(y-y')^2 = 81$ ،

فإن معامل ارتباط بيرسون يساوي r_{xy} ؟

- أ- (27-) ب- (27-) ج- (27-) د- (27-)
36 81 9 54
- 3- احدى القيم التالية يمثل معامل ارتباط سلمي عال جداً:
- أ- (0.7-) ب- (0) ج- (0.98-) د- (0.99)
- 4- إذا كانت قيم X كلها 5,5,5,5,5... وقيم Y كلها 3-,3-,3-,3-... فإن معامل ارتباط سبيرمان=
- أ- 1 ب- (1-) ج- 0 د- 0.12
- 5- إذا أردنا قياس قوة العلاقة واتجاهها بين متغيرين دون البحث في العلاقة السببية فإن الأسلوب الإحصائي المناسب هو:
- أ- تحليل الارتباط ب- تحليل الانحدار ج- التوقع الرياضي د- احتمال الحدث
- 6- إذا أعطيت بيانات عن متوسط درجات الحرارة وكمية الاستهلاك من الكهرباء خلال 50 يوم من أيام الصيف فإنه يمكن القول أن اتجاه العلاقة بين المتغيرين:
- أ- عكسي ب- طردي ج- قوي د- ضعيف
- 7- نوع الارتباط بين مساحة المربع وأبعاده هو:
- أ- تام، بسيط ب- غير تام، بسيط
ج- غير تام، متعدد د- تام، متعدد
- 8- عند حساب معامل الارتباط بين لون العيون ل (50) طالبة ونسبة الذكاء كان معامل الارتباط (0.9) فإنه يمكن تفسير هذا الارتباط القوي ل:
- أ- التأثير بمتغير ثالث فقط ب- التأثير بأكثر من متغير
ج- العلاقة التبادلية د- عامل الصدفة
- 9- القيمة التي تمثل أقوى معامل ارتباط عكسي مما يلي هو:
- أ- (2-) ب- (0.7-) ج- (0.90-) د- (1.5-)

10- قام طالب بحساب معامل ارتباط بيرسون بين متغيرين فوجد أنه $= (1.3)$ فإن ذلك يدل على:

- أ- خطأ في الحساب
ب- عدم وجود ارتباط
ج- طردي تام
د- عكسي تام

11- إذا كان مجموع مربعات فروق الرتب بين (6 قيم) للمتغيرين X, Y هو (50)، فإن معامل ارتباط سبيرمان للرتب بين X, Y يساوي:

- أ- (1.43) ب- (0.43) ج- (0.57) د- (0.43)

12- إذا كان $\Sigma(x-x') = 5$ ، وكان $\Sigma(x-x')^2 = 9$ ، وكان $\Sigma(y-y')^2 = 4$ ، فإن معامل ارتباط بيرسون يساوي $r_{xy} = ?$

- أ- $36/(5-)$ ب- $4/(5-)$ ج- $6/(5-)$ د- $9/(5-)$

13- حسب معامل ارتباط سبيرمان بين X, Y فكان 0.6 فإذا كانت $n=10$ ، فإن Σd^2 يساوي:

- أ- 44 ب- 66 ج- 0.4 د- 0.66

14- إذا كانت $r_1=0.6, r_2=0, r_3=0.43, r_4=-0.9$ فإن معامل الارتباط الذي يعبر عن أقوى علاقة هو:

- أ- r_1 ب- r_2 ج- r_3 د- r_4

س5: ضع رمز الإجابة الصحيحة في الربع المخصص لذلك.

الرقم	1	2	3	4	5	6	7	8	9	10
إجابة										

1- إذا كان $Y = 70 + 0.2 X$ هو خط انحدار Y على X وكانت $Y = 72$ فإن $X = \dots$

- أ- 0.2 ب- 2 ج- 10 د- 20

2- إذا كان $S_x = 0.4, S_y = 0.2$ وكان $Y^* = 2 + 0.3 X$ ، ما قيمة معامل الارتباط بين Y, X ؟

أ- 0.3 ب- 0.4 ج- 0.5 د- 0.6

3- إذا كانت معادلة خط الانحدار النفقات على الدخل (X) هي $Y = 0.3X + 24$ فإذا كان الدخل 1000 دينار فإن حجم النفقات المتوقع هو:

أ- 324 ب- 276 ج- 1024 د- 976

4- إذا أردنا قياس قوة العلاقة واتجاهها بين متغيرين دون البحث في العلاقة السببية فإن الأسلوب الإحصائي المناسب هو:

أ- تحليل الارتباط ب- تحليل الانحدار ج- التوقع الرياضي د- احتمال الحدث

5- حسب الوسط الحسابي للمتغير X فكان 60، وحسب الوسط الحسابي للمتغير Y فكان 70 فإذا كانت معادلة خط الانحدار Y على X هي $Y = 0.5X + b$ فإن قيمة b هي:

أ- 30 ب- 40 ج- 60 د- 70

* إذا كان معامل الارتباط $r_{xy} = 0.90$ ، وأن $S_x = 5.2$ ، $S_y = 4.3$ ، أوجد معادلة الانحدار y/X علماً بأن قيم $x' = 12$ ، $y' = 15$ للإجابة عن الأسئلة 6 - 8

6- أن قيمة b هي: أ- 0.744 ب- 6.07 ج- 0.77 د- 6.75

7- أن قيمة a هي: أ- 0.744 ب- 6.07 ج- 0.77 د- 6.75

8- أن معادلة الانحدار $Y = a + bx$ هي:

أ- $0.744 + 6.07X$ ب- $0.744 + 6.07$

ج- $6.07 + 0.744X$ د- $6.07 + 0.744$

* إذا كان معامل الارتباط $r = 0.952$ ، وأن $S_x = 5.521$ ، وأن $S_y = 5.062$ ، أوجد معادلة انحدار y/x علماً بأن $x' = 13$ ، $y' = 16.571$ اجب عن الأسئلة 9 - 10

9- إن قيمة a تساوي: أ- 0.952 ب- 0.873 ج- 52.2 د- 5.22

10- إن قيمة b تساوي: أ- 0.952 ب- 5.22 ج- 52.2 د- 0.873

الفصل الثاني

التوزيعات المجتمعية الاحتمالية

والتوزيعات العينية

(التوزيع الطبيعي Z ، التوزيع التائي T ،

التوزيع الكائي χ^2 ، التوزيع الفائي F)

Probability Distributions & Sampling Distributions

Introduction	1-2 مقدمة
Normal Distribution	2-2 التوزيع الطبيعي
T Distribution	3-2 التوزيع التائي
χ^2 Distribution	4-2 التوزيع الكائي
F Distribution	5-2 التوزيع الفائي
Sampling Theory	6-2 نظرية المعاينة
Samples	7-2 العينات
Sampling Distributions	8-2 توزيع المعاينة
Exercise	9-2 أسئلة وتمارين

الفصل الثاني

التوزيعات المجتمعية الاحتمالية والتوزيعات العينية

Probability Distributions & Sampling Distributions

1-2 مقدمة Introduction

المتغيرات العشوائية Random Variables: أي متغير عشوائي ينتمي إلى عائلة من التوزيعات، هذه العائلة إما معروفة أو غير معروفة.

المتغيرات العشوائية Random Variables: دالة تمثل العلاقة بين فضاء العينة Ω ومجموعة الأعداد الحقيقية R ولها صفات وخصائص محددة.

المتغير العشوائي المنفصل Discrete Random Variables: يأخذ قيم متميزة ويأخذ عدداً محدوداً ومعدوداً من القيم. مثل عدد الطلاب، عدد السيارات.

المتغير العشوائي المتصل Continues Random Variables: يأخذ مجالاً أو حيزاً على خط الأعداد الحقيقية، مثل الطول، الوزن، درجة الحرارة.

* التوزيعات الطبيعية

1- عائلة التوزيعات الطبيعية Family of Normal Distributions: أفراد هذه العائلة يتحددوا بمعلمتين هما (الوسط الحسابي، التباين) $N(\mu, \sigma^2)$

2- عائلة توزيعات ذات الحدين Binomial Distributions: أفراد هذه العائلة يتحددوا بمعلمتين هما: عدد المشاهدات، احتمال حدوث النجاح. (عدد المشاهدات n ، احتمال حدوث النجاح P) $N(n, p)$

3- عائلة التوزيعات متعددة الحدود Multinomial Distributions: في هذه التوزيعات هناك معالم $(n, p_1, p_2, \dots, p_k)$ ، وهو توسيع لذات الحدين.

4- توزيع بويسون Poison Distributions: له معلمة رئيسية واحدة هي m .

5- التوزيع الاحتمالي لمربعات كاي χ^2 Probability Distribution of χ^2 : وشكل

توزيعه يعتمد على درجات الحرية df

6- التوزيع الاحتمالي الفائي F Probability Distribution F : يعتمد على

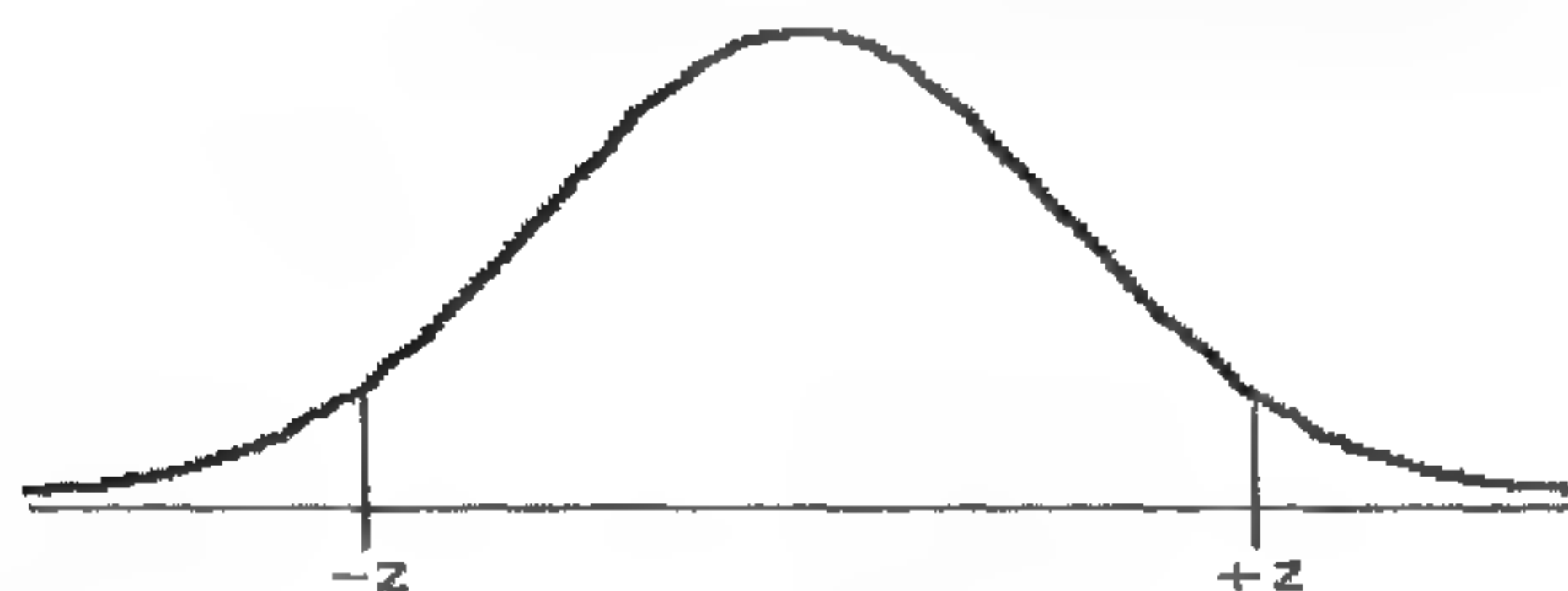
قيمتين لدرجات الحرية $V_1 = n_1 - 1$ و $V_2 = n_2 - 1$

7- التوزيع الاحتمالي التائي T Probability Distribution T : باستخدام

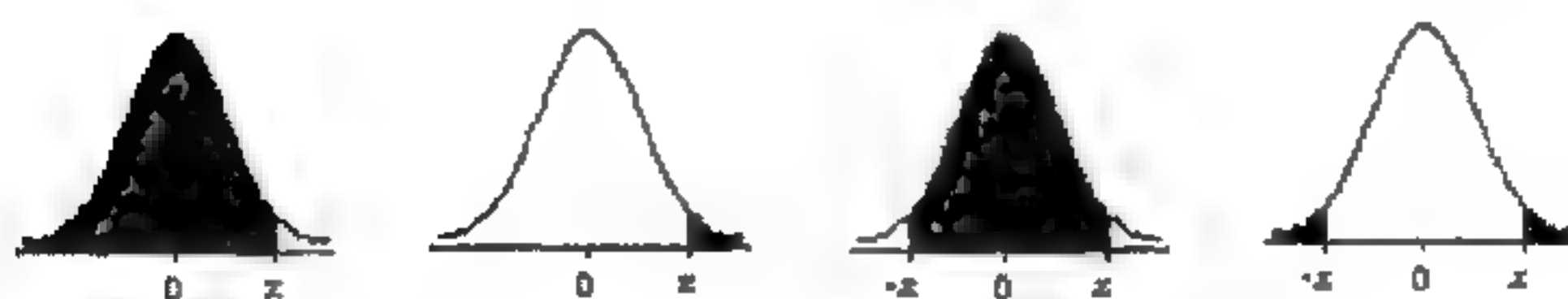
$S_x = S/\sqrt{n}$ هو تقدير الخطأ المعياري $(x_i - \mu)/S$ حيث أن x_i

2-2 التوزيع الطبيعي Normal Distribution

تمثيل البيانات الإحصائية على صورة منحنى.



التوزيع الطبيعي المعياري



☐

☐

☐

☐

Left tail

Right tail

Middle

Two-tail

شكل (2): منحنى التوزيع الطبيعي.

كثير من الباحثين يؤكدون أنه إذا تم أخذ عدد كبير من المشاهدات على إحدى الظواهر مثل الدكاء أو التحصيل الدراسي فإن القيم التي تأخذها هذه المشاهدات تتوزع فيما بينها توزيعاً طبيعياً.

وعلى الرغم مما يحيط بهذا الاصطلاح "التوزيع السوي" من غموض وإبهام فلا بد من التعرف على هذا المفهوم بشكل صحيح.

إن المنحنى السوي هو التمثيل البياني للتوزيع السوي، وهو حالة خاصة من حالات المنحنى الممهد المنتظم وهو ممهد ومنتظم تماماً وقائم على عدد كبير من الحالات، فإنه يمكن الحصول عليه عن طريق التوزيعات التكرارية المتضمنه بيانات حقيقية وليس عن طريق بيانات افتراضية عن طريق الظواهر ذات الصلة.

خصائص التوزيع الطبيعي:

1. خاصية الاستمرارية **Continuity**: عند تمثيل متغير بشكل بياني بمنحنى تكراري فإنه يمكن أن يأخذ أي قيمة ضمن المدى الصفري للتوزيع.

2. خاصية التقارب **Asymptotic**: عند تمثيل متغير بشكل بياني بمضلع تكراري فإنه يمكن الاصطلاح على فئة صفرية عند كل من طرفي التوزيع ويمكن الحصول على مضلع مغلق بمعنى أن المنحنى يكون تقاربياً.

3. خاصية التماثل **Symmetric**: إن نسبة عالية من الأفراد يتجمعون في منطقة المتوسطات ويتوزع الباقي بنسب متفاوتة على جانبي تلك المنطقة، وهذا يعني وجود نقطة انعكاس للمنحنى ويكون ميله عند تلك النقطة يساوي صفراً.

4. صيغة دالته الاحتمالية:

$$F(x) = (1/\sqrt{2\pi\sigma^2}) \text{Exp} [-0.5((x-\mu)/\sigma)^2]$$

ويرمز لهذا التوزيع بالرمز $X \sim N(\mu, \sigma^2)$

5. المساحة تحت المنحنى الطبيعي = 1، وتقسم المساحة إلى نصفين متساويين عند قيمة \bar{X} ، ولإيجاد الاحتمالات الخاصة بالتوزيع الطبيعي نستخدم المعادلة:

$$P(a < x \leq b) = \int_a^b f(x)dx$$

العلامات المحولة **Transformed Scores**

إن المشاهدة الخام غير قابلة للتفسير لأنها ليست كافية لإعطاء صورة عن الوضع الصحيح للمشاهدة، لأنه لا بد من معرفة الوسط الحسابي، وأعلى وأدنى مشاهدة، وأي إحصائي تشتت عن هذه المشاهدات.

أنواع التحويلات:

تصنف الدرجات المحولة تحت نوعين من التحويلات هما:

1- التحويل الخطي Linear

- تحويل المشاهدات من مقياس لآخر بحيث لا يحدث أي تغيير على شكل التوزيع.
- تكون العلاقة بين المشاهدات على المقياسين خطية.
- من أمثله تحويل العلامة الخام إلى علامة زائية Z score.

2- التحويل غير الخطي Non Linear

- يصاحب تغير الدرجات تغير في شكل التوزيع الأصلي.
- تكون العلاقة بين المشاهدات على المقياسين ليست خطية.
- من أمثله تحويل العلامة الخام إلى علامة معيارية زائية طبيعية Z_{score}.

* الكشف عن نوع التحويل:

- رسم شكل الانتشار.
- - تمثيل المشاهدات بيانياً.
- إيجاد الفروق بين المشاهدات المتتالية على المقياسين.

* أمثلة على التحويل الخطي:

1- العلامة المعيارية الزائية (Z - Score)

علامة معيارية في توزيع وسطه الحسابي = 0 ، وانحرافه المعياري = 1 ، يتم الحصول عليه بالتحويل الخطي دون أي تغيير في شكل التوزيع الأصلي.

العلامة المعيارية = $\frac{\text{العلامة الخام} - \text{الوسط الحسابي}}{\text{الانحراف المعياري}}$

$$Z = \frac{X - \bar{X}}{\sigma_x}$$

X : المشاهدة. \bar{X} : الوسط الحسابي. σ_x : الانحراف المعياري.

* خصائص العلامات الزائية:

- كل مشاهدة يمكن تحويلها إلى مشاهدة زائية.
- الوسط الحسابي للملاحظات الزائية = 0، والانحراف المعياري = 1
- تحافظ على شكل التوزيع.

2- العلامة المعيارية التائية (T - Score)

لستخلص من الخصائص غير المرغوبة للعلامات الزائية فإنه يمكن تحويلها إلى أخرى بوسط حسابي وانحراف معياري جديدين، ومن أشهر هذه التحويلات هي العلامة المعيارية التائية بوسط حسابي = 50، وانحراف معياري = 10 كما تبينه العلاقة أدناه:

$$T = 10 * Z + 50$$

مثال) إذا كانت العلامة المعيارية الزائية = 1، فإن العلامة المعيارية التائية = 60، أما إذا كان التوزيع محدد وسطه الحسابي وانحرافه المعياري مسبقاً، فإن معادلة التحويل تكون:

$$Z' = Z \sigma + \mu$$

* أمثلة على التحويل غير الخطي:

1- العلامة الزائية المعدلة (Normalized Z - Score)

علامة معيارية في توزيع وسطه الحسابي = 0 ، وانحرافه المعياري = 1 ، حول شكله الأصلي إلى الشكل الطبيعي.
خطوات التحويل:

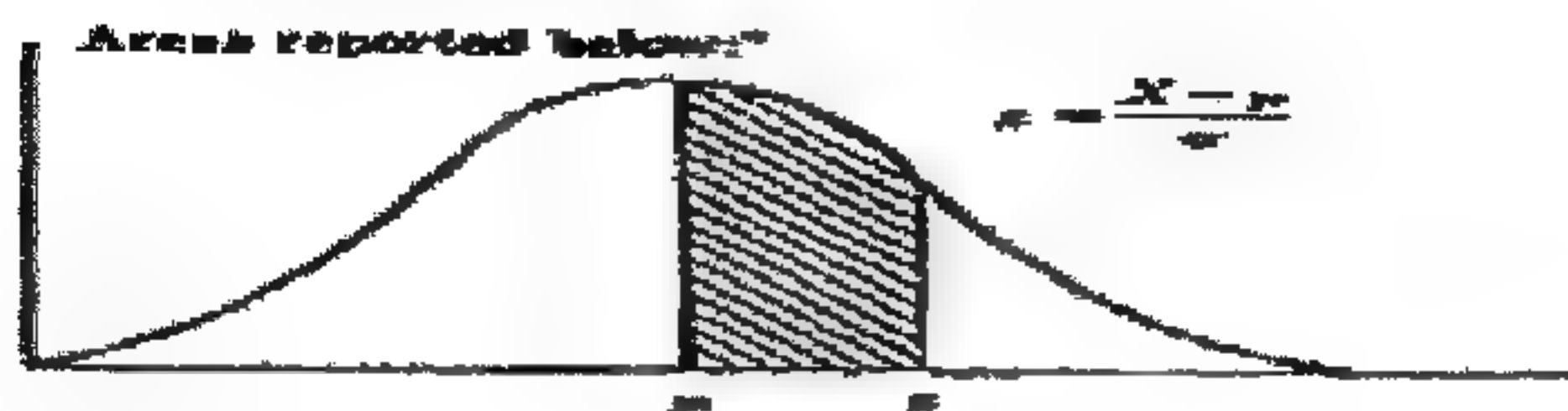
تحويل كل علامة خام في التوزيع الأصلي إلى رتبة مثنية.
إيجاد العلامة المعيارية الزائية المقابلة لكل رتبة مثنية من جدول التوزيع الطبيعي.

2- علامة ستانين (التساعي) (Stanine Score)

علامة معيارية في توزيع وسطه الحسابي = 5 ، وانحرافه المعياري = 2 ، وأقل علامة في التوزيع هي ستانين 1 وأعلى علامة هي ستانين 9 وكل ستانين يقابل نسبة معينة من العلامات في التوزيع الأصلي.

التوزيع الطبيعي Normal Distribution

الجدول أدناه يعطي المساحة تحت منحنى التوزيع الطبيعي القياسي (المساحة ما بين الوسط وقيمة Z)



وقد تم تحويل قيم التوزيع الطبيعي إلى التوزيع الطبيعي المعياري، وهو توزيع له متوسط حسابي $\mu = 0$ ، وانحراف معياري $\sigma = 1$ ، ويرمز للمتغير العشوائي المعياري بالحرف Z. إن التوزيع الطبيعي أكثر التوزيعات الاحتمالية استخداماً لأن توزيع المتغيرات يكون طبيعياً في أكثر الحالات التطبيقية، كما أنه يمثل تقديراً دقيقاً لعدد كبير من التوزيعات الأخرى إذا كان عدد المتغيرات كبيراً، وأن لكل توزيع معالم وللتوزيع الطبيعي معلمتان هما الوسط الحسابي للمجتمع μ والانحراف المعياري للمجتمع σ ، لذلك عرف التوزيع الطبيعي المعياري بأنه توزيع له متوسط حسابي $\mu = 0$ ، وانحراف معياري $\sigma = 1$ ، ولكثرة استخدام التوزيع الطبيعي المعياري فقد دون في جداول كما هو موضح في (ملحق رقم 1) ليسهل استخدامه في الحالات التطبيقية، وحدود هذه الجداول هي من $(-3$ إلى $+3)$ ، وإن المساحة الكلية للتوزيع $\mu = 1$ والمساحة فوق Z تساوي صفر $= 0.5$ ، والمساحة تحت Z تساوي صفر $= 0.5$.

العلامة المعيارية الزائفة (Z - Score)

علامة معيارية في توزيع وسطه الحسابي $= 0$ ، وانحرافه المعياري $= 1$ ، يتم الحصول عليه بالتحويل الخطي دون أي تغيير في شكل التوزيع الأصلي.

$$\text{العلامة المعيارية} = \frac{\text{العلامة الخام} - \text{الوسط الحسابي}}{\text{الانحراف المعياري}}$$

الانحراف المعياري

$$Z = \frac{X_i - \mu}{\sigma_x}$$

$$Z = \frac{X - \bar{X}}{\sigma_x}$$

حيث:

X_i : المشاهدات.

\bar{X} : الوسط الحسابي.

μ : الوسط الحسابي للمجتمع الذي قيمه X_i .

σ_x : الانحراف المعياري للمجتمع الذي قيمه X_i .

* خصائص العلامات الزائفة:

- كل مشاهدة يمكن تحويلها إلى مشاهدة زائفة.
- الوسط الحسابي للمشاهدات الزائفة = 0، والانحراف المعياري = 1
- تحافظ على شكل التوزيع.

خواص المنحنى الطبيعي Normal Distribution Curve

- المساحة الكلية تحت المنحنى = 1
- شكل المنحنى على هيئة جرس.
- تتركز المشاهدات حول الوسط الحسابي.
- المنحنى متماثل حول الوسط الحسابي.
- إن المساحة الموجودة على بعد
- ± 1 انحراف معياري = 68.27% عن \bar{X}
- ± 2 انحراف معياري = 95.45% عن \bar{X}
- ± 3 انحراف معياري = 99.73% عن \bar{X}

99-2003- إذا كانت قيمة الوسط الحسابي والوسيط والمنوال متماثلة تماماً في جدول

تكراري فإن البيانات يكون لها توزيع يسمى:

أ) متماثلاً (طبيعياً) ب) مائلاً إلى اليمين ج) مائلاً إلى اليسار د) مقعراً.

من خصائص التوزيع الطبيعي أن الوسط = الوسيط = المنوال.

مثال 2-1) إذا كانت علامة طالب في مادة الإحصاء هي 75 ومتوسط علامات الإحصاء هو 90 والانحراف المعياري لعلامات الإحصاء هو 5 وعلامته في مادة النظم هي 70 ومتوسط علامات النظم هو 60 والانحراف المعياري لعلامات النظم هو 4 فأأي العلامات أفضل.

$$Z_{stat} = \frac{X - \bar{X}}{\sigma_x} = \frac{75 - 90}{5} = \frac{-15}{5} = -3$$

$$Z_{Mis} = \frac{X - \bar{X}}{\sigma_x} = \frac{70 - 60}{4} = \frac{10}{4} = 2.5$$

الذي يحصل على علامة زائية أعلى هو الأفضل.

إذا علامته في مادة النظم MIS أفضل من علامته في مادة الإحصاء.

80-2005- إذا كان الوسط الحسابي لعلامات طلبة احد الصفوف في مادة الإحصاء

(60)، بانحراف معياري مقداره (5)، وكانت علامة أحد الطلبة في هذه المادة (75)، فإن

قيمة العلامة المعيارية لهذا الطالب هي:

أ- (3-) ب- 3/1 ج- $\sqrt{3}$ د- 3

$$\text{العلامة المعيارية} = \frac{\text{العلامة الخام} - \text{الوسط الحسابي}}{\text{الانحراف المعياري}} = \frac{60 - 75}{5} = \frac{-15}{5} = -3$$

* المساحة تحت المنحنى الطبيعي

إن القيمة الاحتمالية لـ Z يتم إيجادها من الجداول الإحصائية للتوزيع الطبيعي، فمثلاً لإيجاد احتمال القيمة 0.3413 نقوم بتحديد موقع القيمة في الجهة اليمين من الجدول ثم نجد القيمة المقابلة لها في الجانب الأيسر وهي 1 وهكذا.

مثال: إذا كان X يمثل توزيعاً طبيعياً بمتوسط $\mu=60$ وانحراف معياري $\sigma=10$

والمطلوب إيجاد الاحتمال التالي:

$$1. P(60 < X \leq 65) \quad 2. P(65 \leq X \leq 70) \quad 3. P(55 \leq X \leq 60)$$

الحل: 1- $P(60 \leq X \leq 65)$

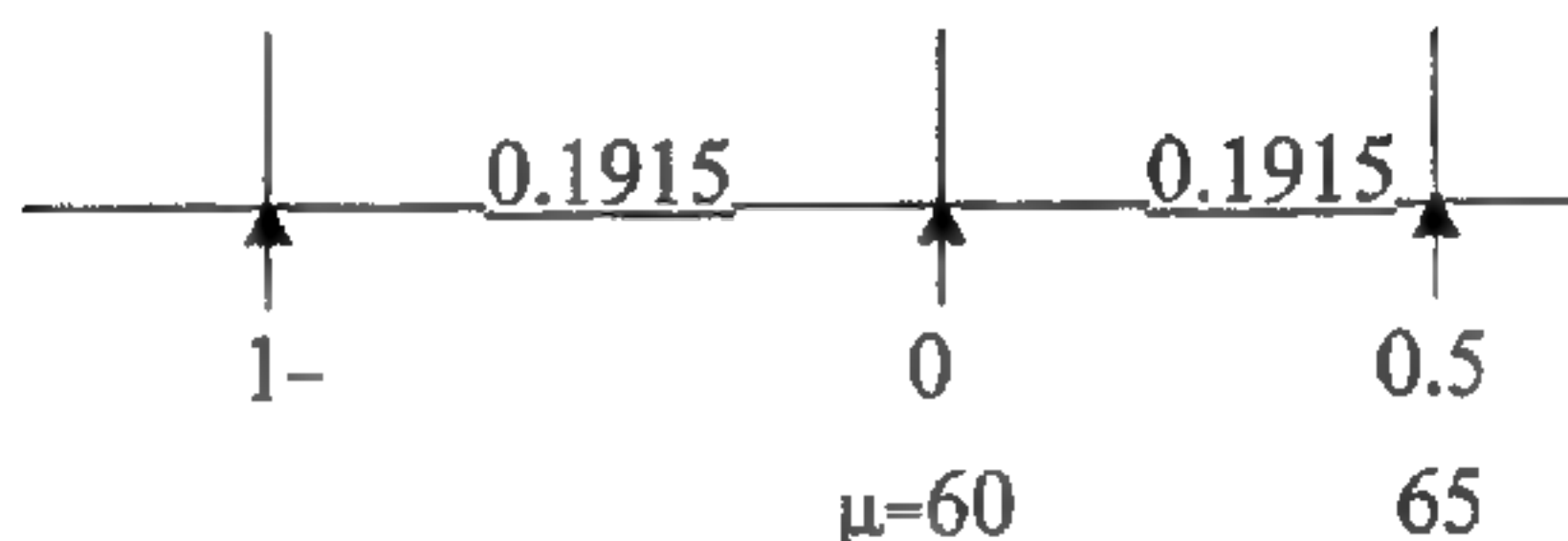
نقوم بتحويل قيم X للتوزيع الطبيعي إلى قيم Z المعيارية باستخدام الصيغة $Z = \frac{X - \mu}{\sigma}$

قيمة Z عند $X = 60$

$$Z_1 = \frac{X - \mu}{\sigma} = \frac{60 - 60}{10} = 0$$

قيمة Z عند $X = 65$

$$Z_1 = \frac{X - \mu}{\sigma} = \frac{65 - 60}{10} = 0.5$$



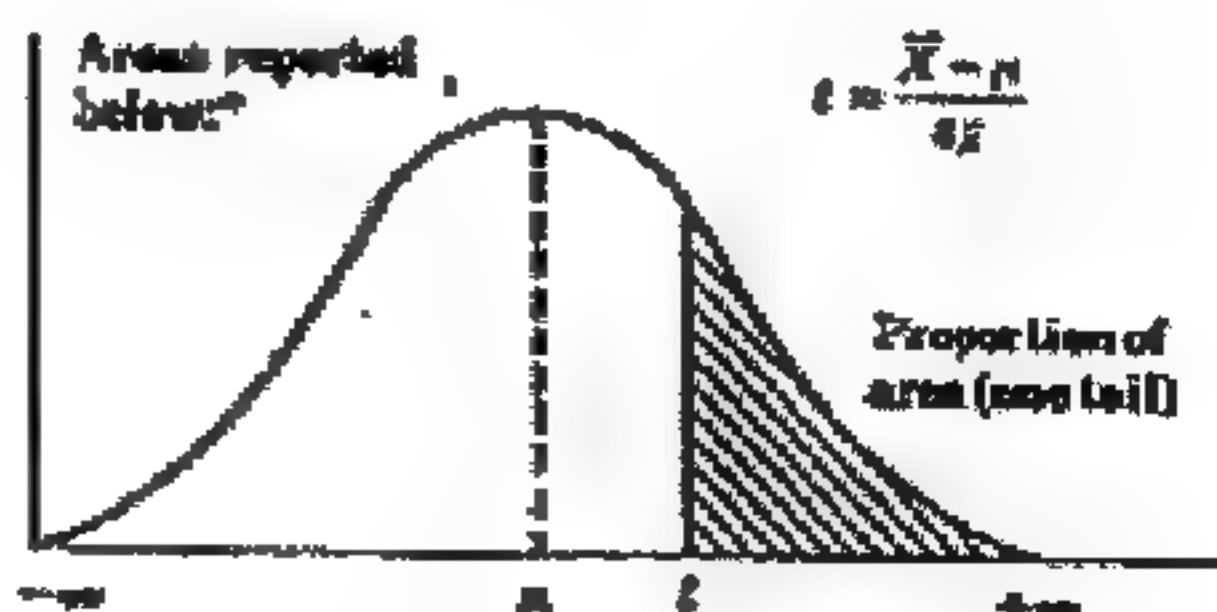
وبالرجوع لجدول التوزيع الطبيعي نجد أن قيمة الاحتمال بين 0-0.5 هي 0.1915

3-2 التوزيع الاحتمالي التائي Probability Distribution (T)

الجدول أدناه يعطي قيمة t

المقابلة للمساحة المظلمة وقيمها α

Proportions of Area for the t Distributions



في حالة التوزيع الطبيعي Z إن الانحرافات للمتوسطات العينية $(\bar{x}_i - \mu)$ وقسمتها على الخطأ المعياري σ_x موزعة توزيعاً طبيعياً مع $\mu=0$ و $\sigma=1$ ، حيث أن طرح μ من وسط كل

عينة \bar{x}_i لا يغير من شكل التوزيع لمتوسطات العينة وأن قسمتها على $\sigma_{\bar{x}}$ يؤدي إلى تقليل التباين إلى 1، ولكن عند حساب الانحراف المعياري لكل متوسط \bar{x}_i باستخدام $(\bar{x}_i - \mu)/S_{\bar{x}_i}$ حيث أن $S_{\bar{x}_i}$ هو تقدير الخطأ المعياري لمتوسط العينة \bar{x}_i ، يجعل توزيع الانحرافات أكثر وأوسع تسطحاً مما هي عليه في حالة Z ، والسبب في هذا أن المقام هو الخطأ المعياري لعينة وليس للمجتمع، وأن هذه الزيادة في الانحراف تكون انعكاساً للتباين الكبير للنسبة $(\bar{x}_i - \mu)/S_{\bar{x}_i}$ ، وأن التوزيع المتوقع لهذه النسبة هو ما يسمى (توزيع t) وصيغة احتسابه كالتالي:

$$t = (\bar{x} - \mu)/S_{\bar{x}}$$

$$S_{\bar{x}} = S/\sqrt{n}$$

حيث أن:

شكل العام لتوزيع t لعدد من درجات الحرية



شكل (3): الشكل العام لتوزيع t لعدد من درجات الحرية.

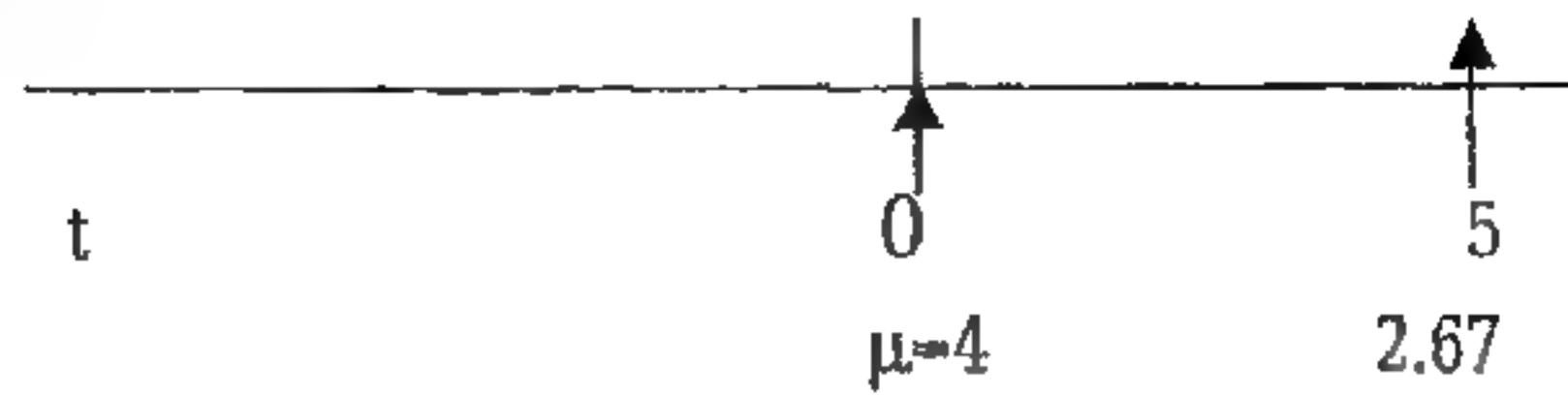
إن توزيع t له نفس خصائص التوزيع الطبيعي Z حيث أنه: متماثل، يمتد من $-\infty$ إلى $+\infty$ يختلف عن التوزيع الطبيعي من كون شكله يعتمد على عدد درجات الحرية V حيث أن $V=n-1$ حيث أن n هي حجم العينة المعتمدة في حساب التباين ويرمز لدرجات الحرية بالرمز df وغالباً ما يقترب توزيع t من شكل التوزيع الطبيعي كلما زادت عدد درجات الحرية وأكثر درجة يتقارب التوزيعان بها عندما تكون $df=30$ ، وعندما تكون $df=\infty$ يصبح توزيع t توزيعاً طبيعياً، لذلك هناك جدول خاص بتوزيع t يعطي قيمه بناء على قيم df ويوجد هذا الجدول بملاحق الكتاب.

مثال: إذا كان لدينا عينة حجمها $n=16$ وانحرافها المعياري $S=1.5$ مسحوبة من مجتمع

يخضع لتوزيع t ، وسطه الحسابي $\mu=4$ فما هو احتمال أن تكون قيمة الوسط $\bar{x}=5$

$$S_{\bar{x}} = S/\sqrt{n} = 1.5/\sqrt{16} = 1.5/4 = 0.375$$

$$t = \frac{\bar{x} - \mu}{S_x} = \frac{5-4}{0.375} = \frac{1}{0.375} = 2.67$$



ومن جدول t وبدرجة حرية $15 = 16-1 = V=n-1$ والبحث عن اقرب قيمة لـ 2.67 المحسوبة نجدها 0.99 وعليه فإن $P(t=2.67) = 0.99$

4-2 التوزيع الاحتمالي لمربعات كاي Probability Distribution of χ^2

Chi-Square Distribution

توزيع كاي تربيع



شكل (4): توزيع كاي تربيع

توزيع χ^2 من التوزيعات الاحتمالية النظرية المستمرة، ودالة كثافته النوعية تأخذ القيم من 0 وحتى $+\infty$ ، ولكن شكله لا يشبه توزيع Z ، لأنه منحني يقترب للتماثل مع التوزيعات من جانبها الأيمن فقط، وشكل توزيعه يعتمد على درجات الحرية df لذلك هو دالة لعدد درجات الحرية ويبدأ بالشكل L ويقترب من التماثل كلما تزداد درجات الحرية.

$$\chi^2 = \frac{(n-1) S^2}{\sigma^2}$$

وهي توزيع مع $(n-1)$ من درجات الحرية.

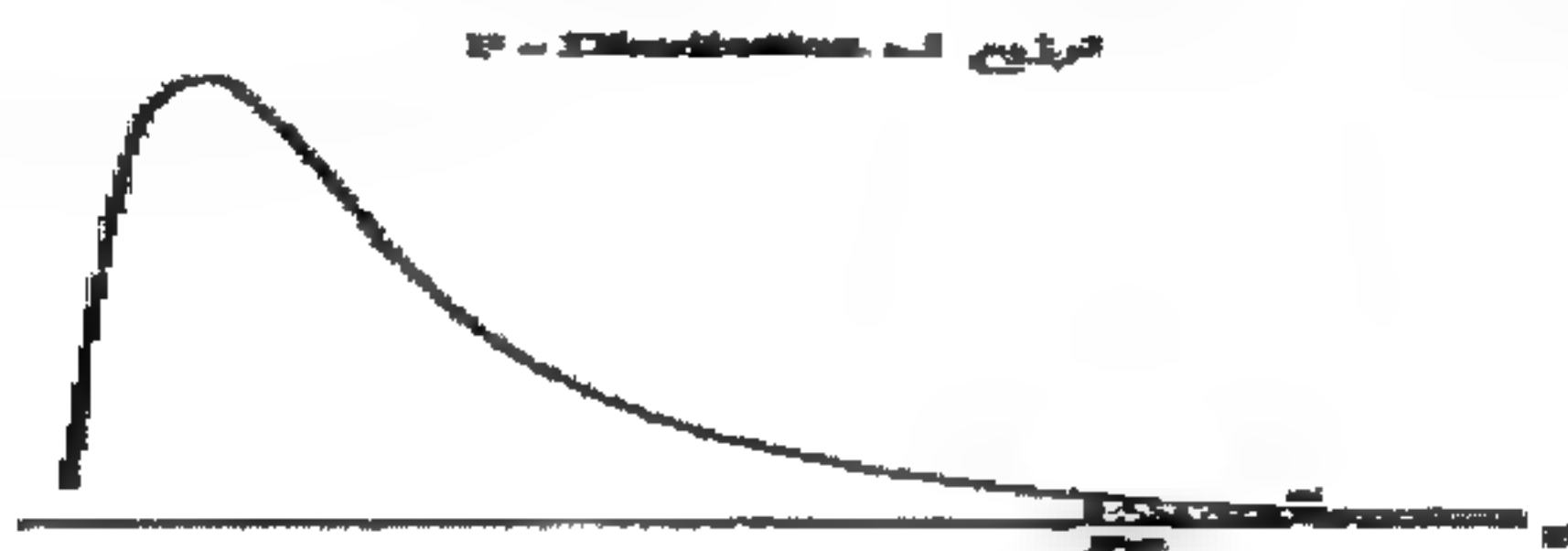
لإيجاد المساحات تحت منحنى χ^2 نستعمل جدول χ^2 حيث يمثل العمود الأيسر درجات الحرية V والخط الأفقي يمثل مستويات المعنوية، والقيم داخل الجدول تمثل المساحات لتوزيع. مثال: إذا كان لدينا χ^2 المحتسبة هي 0.995 لعينة حجمها $n=12$ فما هي قيمة χ^2 التي يقع إلى يسارها 0.995 من المساحة.

الحل:

$$V = n - 1 = 12 - 1 = 11 \quad \text{من الجدول } \chi^2_{0.995,11} = 26.757$$

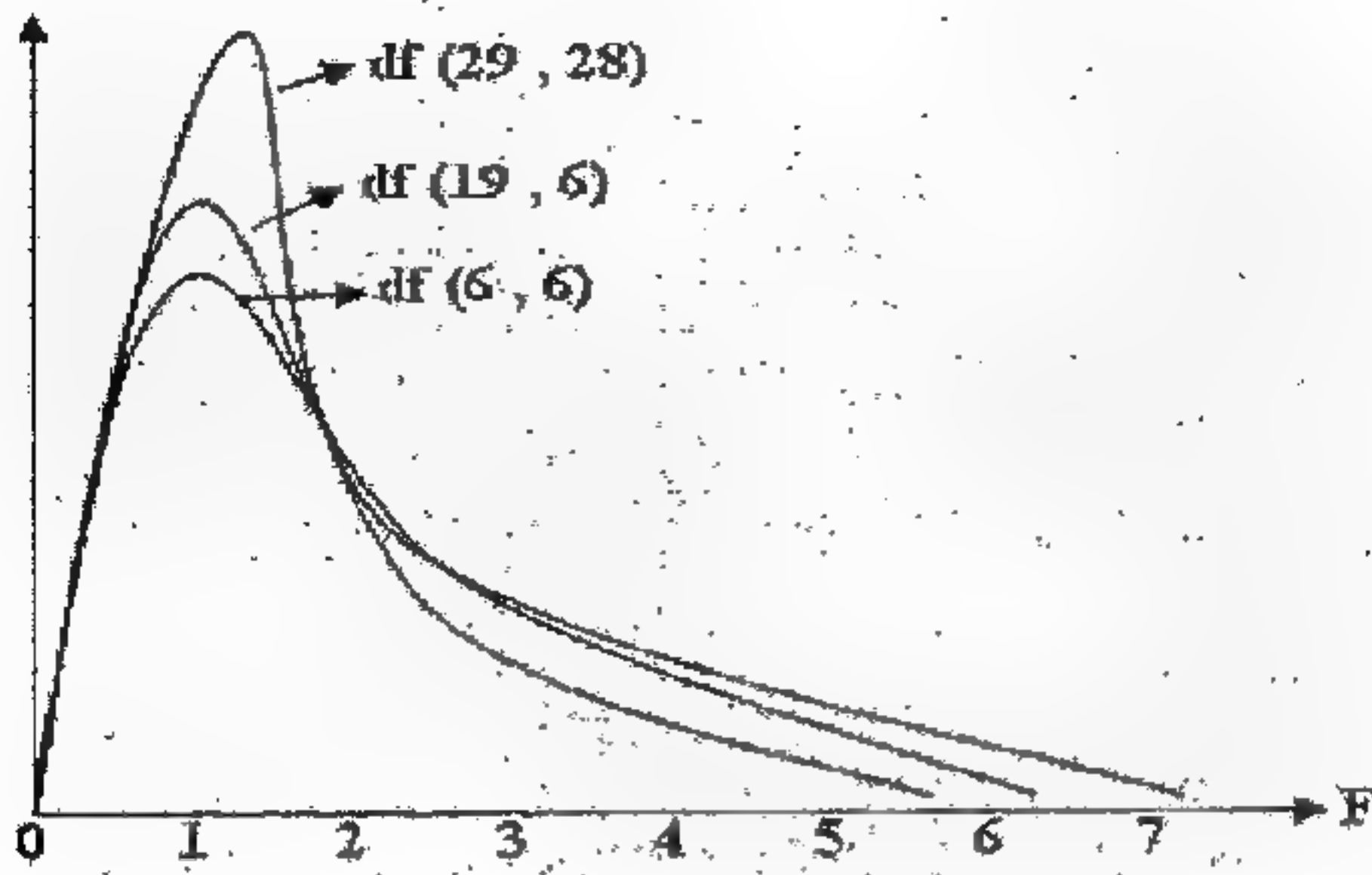
5-2 التوزيع الاحتمالي الفائي F

F- Probability Distribution



شكل (5): توزيع ف F

توزيع F : هو توزيع ملو جهة اليمين بمعلمتين تتمثلان بدرجة حرية (البسط ، المقام) وهما $k - 1$ للبسط ، $k - n$ للمقام حيث n مجموع أحجام العينات، فإذا كان لدينا اختبار لقياس معنوية الفرق بين التقديرين (F) نوجد F_α حيث α مستوى المعنوية المستخدم للفرضية H_0 التي ترفض إذا كان $F < P_\alpha$ وإلا نؤكد بوجود الاختلاف بين المتوسطات، والشكل التالي يبين توزيع F



شكل (6): منحنى توزيع F حسب درجات الحرية

كما في توزيعات χ^2 ، t فإن توزيع F هو من التوزيعات الاحتمالية النظرية المستمرة، ولكن شكله لا يشبه توزيعي χ^2 ، t لأنه يعتمد على قيمتين لدرجات الحرية $V_1=n_1-1$ و $V_2=n_2-1$ وكل منهما تبدأ من 1 وحتى ∞ ، ونسبة التباين تعتمد على تباين العينتين S^2_1/S^2_2 ، ونظرياً فإن هذه القيمة تقترب من 1 لأن التباينين هما تقديرات لنفس الكمية ويرمز لها بالرمز F_s .

إذا كانت العينتين من مجتمعين طبيعيين منفصلين مختلفين في المتوسطات ولها نفس التباين أي: $\sigma^2_1=\sigma^2_2$ ، $\mu_1\neq\mu_2$ أو أن العينتين مسحوبة من نفس التوزيع الطبيعي ولكن تباينهما نفسه، يرمز للتوزيع بالرمز F .

والشكل العام لتوزيع F كما هو شكل توزيع χ^2 يبدأ على شكل حرف L عند درجات الحرية الصغيرة، ثم يصبح مفرطح باتجاه اليمين بازدياد درجات الحرية.

يستخدم الجدول في ملحق الكتاب لإيجاد المساحات تحت المنحنى لتوزيع F

مثال: جد المساحة تحت المنحنى لتوزيع F عند النقطة 0.95 ودرجات حرية $v_1=10$

و $v_2=8$.



ومن جدول F وبدرجتي حرية 8, 10 حيث $v_1=10$ بالاتجاه الأفقي للجدول و $v_2=8$ بالاتجاه العمودي للجدول و $P=0.95$ وبذيل واحد نجد القيمة 3.347 فتكون المساحة:

$$F = \frac{1}{F_{(0.95,10,8)}} = \frac{1}{3.347} = 0.299$$

6-2 نظرية المعاينة Sampling Theory

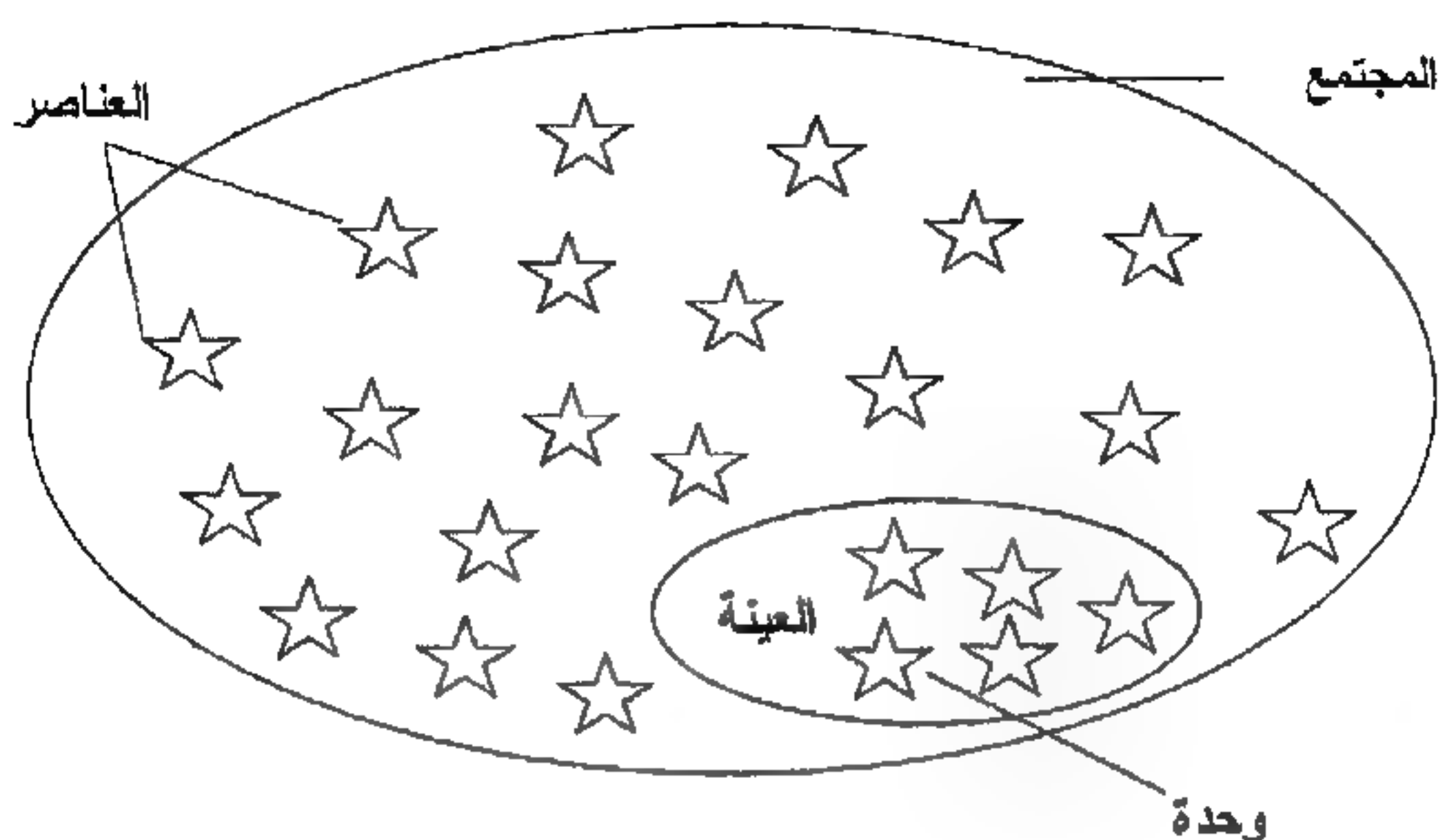
في الإحصاء الاستدلالي حيث يتوصل الباحث الى خصائص المجتمع Population عن طريق العينة Sample، أو المعاينة، لأن دراسة المجتمع احياناً مستحيلة أو صعبة جداً ومكلفة وتحتاج الى وقت وجهد كبيرين.

المجتمع والعينة والعنصر ووحدة المعاينة

Population, Sample, Element, and Sampling Unit.

إن أسئلة البحث والأهداف التي نحتاجها هي التي تحدد مدى الحاجة إلى استخدام العينة، كما أن قيود الوقت والكلفة وطبيعة المجتمع المبحوث قد تحول دون قدرة الوصول إلى المجتمع الكامل لجمع البيانات منه، وهنا لابد من إتباع أحد أساليب المعاينة والتي تزودك بأساليب مختلفة لأنواع العينات.

والشكل التالي يبين العلاقة بين المجتمع والعينة والعنصر ووحدة المعاينة.



الشكل (1): المجتمع والعينة والعنصر ووحدة المعاينة.

Source: Saunders, Mark, Lewis, Philip, & Thornhill, Adrian (2007). *Research methods for business students* (4th ed.). Edinburgh Gate, Harlow: Pearson Education Limited. p. 205.

التوزيع **Distribution**: مجموعة مشاهدات مهما كان عددها.

المجتمع **Population**: جميع الأفراد التي يمكن أن يأخذها المتغير.

مجتمع العينة **Sample Population**: المجتمع الذي تؤخذ منه العينة.

مجتمع الهدف **Target Population**: المجتمع الذي ستعمم عليه نتائج الدراسة التي

أجريت على مجتمع العينة.

العينة **Sample**: مجموعة جزئية من المجتمع.

المؤشر **Index**: تدل على جميع مقاييس النزعة المركزية والتشتت والعلاقة (الارتباط)

سواء محسوبة لعينات أو لمجتمعات.

مؤشر عينة (إحصائي) **Statistic**: يستخدم للعينات، مثل الوسط الحسابي لعينة X' .

مؤشر مجتمع (معلم) **Parameter**: يستخدم للمجتمع، مثل الوسط الحسابي μ .

7-2 العينات Samples

المجتمع الإحصائي Population: أي تجمع معرف من الأشياء أو الأشخاص أو الحوادث، وهو المجموعة الشاملة التي يجري اختيار العينات منها.

العينة Sample: جزء من المجتمع تتم دراسة الظاهرة عليهم من خلال المعلومات عن هذه العينة، حتى تتمكن من تعميم النتائج على المجتمع.

العينة Sample: أي مجموعة جزئية من المجتمع الإحصائي يتم جمع البيانات من خلالها بصورة مباشرة.

خصائص العينة تختلف باختلاف العينة وحتى يكون التقدير مناسب يجب أن تكون العينة تتمتع بما يلي:

– أن تكون العينة ممثلة للمجتمع.

– أن يكون حجم العينة مناسب.

العينة الممثلة للمجتمع: هي العينة التي يتم اختيارها بطريقة عشوائية.

أسباب استخدام العينات Reasons for using Samples:

1. صعوبة حصر أفراد المجتمعات في وقت واحد.
2. تقليل نفقات الدراسة.
3. صعوبة تأمين العدد الكافي من المختصين الذين تحتاجهم الدراسة.
4. التقليل من الوقت اللازم لإجراء الدراسة والتمكن من تحديده.
5. الحصول على دقة قريبة من استخدامنا للمجتمع.

عدد أفراد العينة Number of Sample's Persons:

1. لا يوجد قانون محدد لتحديد حجم العينة.
2. الدراسات المسحية: 20% من أفراد المجتمع إذا كان صغير نسبياً (500-1000) وتصبح 5% من أفراد المجتمعات الكبيرة جداً.
3. العينة تكون 30 فرداً من أفراد المجتمعات الصغيرة. ولا تقل عن ذلك.
4. الدراسات الارتباطية: 30 فرداً لكل متغير في الارتباط والانحدار المتعددين.

5. البحوث التجريبية: 15 فرداً في كل مجموعة.

6. التحليل العاملي: أن يكون حجم العينة من خمسة إلى عشرة أمثال عدد الفقرات.

ويبين الجدول (1.4) حجم العينة المطلوب اعتماداً على هامش الخطأ المسموح به، ومنه يتبين أنه كلما قل هامش الخطأ المسموح به زاد حجم العينة.

الجدول (1) اختلاف حجم العينة المطلوب باختلاف هامش الخطأ المسموح به.

هامش الخطأ المسموح به (Margin of error)	حجم المجتمع الكلي			
	(%1)	(%2)	(%3)	(%5)
50	50	49	48	44
100	99	96	91	79
150	148	141	132	108
200	196	185	168	132
250	244	226	203	151
300	291	267	234	168
400	384	343	291	196
500	475	414	340	217
750	696	571	440	254
1000	906	702	516	278
2000	1655	1091	696	322
5000	3288	1622	879	357
10000	4899	1936	964	370
100000	8762	2345	1056	383
1000000	9513	2395	1066	384
10000000	9595	2400	1067	384

Source: Saunders, Mark, Lewis, Philip, & Thornhill, Adrian (2007). *Research methods for business students* (4th ed.). Edinburgh Gate, Harlow: Pearson Education Limited. p. 212.

كما قدم كل من كريجيسي ومورجان (Krejcie & Morgan, 1970) جدول يسهل اتخاذ قرار حيّد لتحديد حجم العينة المطلوبة اعتماداً على حجم المجتمع الكلي وهامش خطأ المسموح به (5%).

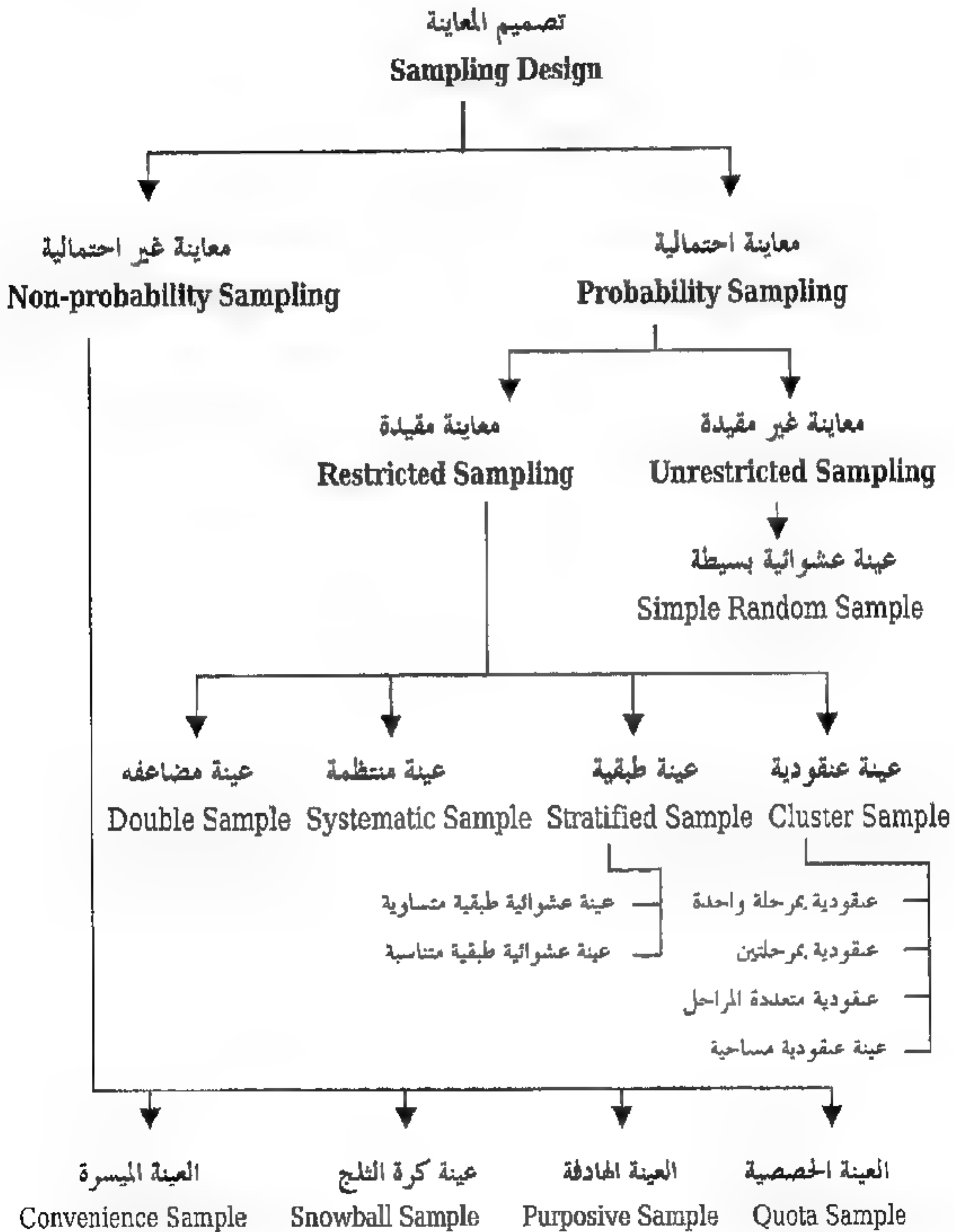
الجدول (2)

تحديد حجم العينة اعتماداً على حجم المجتمع الكلي (هامش الخطأ المسموح به (5%)).

العينة (n)	المجتمع (N)	العينة (n)	المجتمع (N)	العينة (n)	المجتمع (N)
291	1200	140	220	10	10
297	1300	144	230	14	15
302	1400	148	240	19	20
306	1500	152	250	24	25
310	1600	155	260	28	30
313	1700	159	270	32	35
317	1800	162	280	36	40
320	1900	165	290	40	45
322	2000	169	300	44	50
327	2200	175	320	48	55
331	2400	181	340	52	60
335	2600	186	360	56	65
338	2800	191	380	59	70
341	3000	196	400	63	75
346	3500	201	420	66	80
351	4000	205	440	70	85
354	4500	210	460	73	90
357	5000	214	480	76	95
361	6000	217	500	80	100
364	7000	226	550	86	110
367	8000	234	600	92	120
368	9000	242	650	97	130
370	10000	248	700	103	140
375	15000	254	750	108	150
377	20000	260	800	113	160
379	30000	265	850	118	170
380	40000	269	900	123	180
381	50000	274	950	127	190
382	75000	278	1000	132	200
384	1000000	285	1100	136	210

Source: Sekaran, Uma (2003). *Research methods for business: A skill-building approach* (4th ed.). New York: John Wiley & Sons Inc., p. 294.

تصنف العينات إلى عينات احتمالية وعينات غير احتمالية:
يبين الشكل (2) الأساليب المختلفة التي يمكن إتباعها في اختيار العينة.



أولاً: العينات غير الاحتمالية Non probabilistic Samples

هي عينات يتدخل فيها ميل الباحث وتمييزه بدرجة كبيرة في اختيار أفرادها، ويصعب تعميم نتائجها على جميع أفراد المجتمع. ومنها ما يلي:

1. العينة الميسرة Convenience Sample

تتضمن العينة الميسرة اختيار جزائي أو مصادفة للحالات المدروسة والتي من السهولة الحصول عليها في العينة، إذ يتم اختيار وحدات العينة بناء على سهولة الوصول والاتصال بالأعضاء، وهي سريعة التنفيذ وقليلة الكلفة، ولكن لا يمكن تعميم نتائجها. وغالباً ما تخدم هذه العينة كدراسة قبلية/ أولية أكثر من كونها عينة مهيكلية متكاملة (Saunders et al., 2007, 234). ومثال ذلك إذا أراد الباحث أن يتعرف على رأي الطلبة المبدئي في أداء المواصلات في جامعة ما فإنه يقوم بسؤال أول (50) طالب أو طالبة يواجههم عند البوابة الرئيسة ليتعرف على آرائهم في أداء المواصلات في تلك الجامعة.

2. العينة الهادفة Purposive Sample

تستخدم العينة الهادفة للحصول على معلومات من شريحة مُحددة قادرة على توفير المعلومات، إما بسبب موقعهم، أو لأن بعض المعايير التي وضعها الباحث تتوفر فيهم؛ لأنهم أفضل الأشخاص القادرين على توفير المعلومات، حيث يتم اختيار وحدات العينة بناء على الخبرات في الموضوع الذي يدرس. وتستخدم العينة الهادفة عندما تكون المعلومات المطلوبة متوفرة لدى فئة معينة من الأفراد، فهي التي تملك المعرفة في الموضوع المبحوث وتستطيع تقديم المعلومة.

وتستخدم العينة الهادفة في الغالب عندما نتعامل مع عينات صغيرة، أو عندما نتعامل مع حالات نريد منها معلومات خاصة (Neuman, 2000).

ومثال ذلك: لو أردنا الوصول إلى إجابة السؤال التالي: ماذا يلزم المديرية للوصول إلى مراكز القمة؟ فإن العينة الهادفة المناسبة هنا هي مجموعة النساء التي احتلت مراكز عليا؛ إذ يكون لديهن معرفة متخصصة في ذلك الموضوع نتيجة الخبرة.

3. العينة الحصصية Quota Sample

تستخدم العينة الحصصية في مقابلات المعاينة (Survey) وهي غير عشوائية تماماً، وتقوم على افتراض أن العينة تمثل المجتمع وأن التغير بالنسبة لمتغيرات العينة الحصصية هي نفسها بالنسبة لمتغيرات المجتمع، لذا فإن العينة الحصصية هي نوع من العينة العشوائية الطبقية ولكنها تختار أفراد الطبقة بطريقة غير عشوائية، إذ تعتمد على تقسيم المجتمع إلى مجموعات خاصة، ثم حساب حصة كل مجموعة اعتماداً على علاقتها بالبيانات المتوفرة وحجم المجتمع، ثم الحصول على تلك الحصة بأيسر الطرق (Saunders et al., 2007, 226).

وتستخدم العينة الحصصية عندما يكون هناك صفات محدّدة يجب أن تؤخذ مسبقاً بالاعتبار في العينة مثل: (الجنس، الوظيفة، التوزيع الجغرافي)، إذ لا بد والحالة هذه من توزيع العينة بالحصة على المجتمع لتمثل التنوع بداخله.

ومثال ذلك: إذا أردنا توجيه سؤال معين إلى مجموع العمال والعاملات في الشركة المتحدة، وتبين إن العمالة في الشركة تتكون من (30%) من العمال الذكور، بينما نسبة (70%) من العمالة إناث، وتقرر أن يكون إجمالي العينة (10) أشخاص. فإننا سنوجه السؤال إلى أول (3) عمال ذكور، وأول (7) عاملات إناث تتم مواجعتهم في ظروف مريحة وبصورة كيفية دون الاعتماد على الأسلوب العشوائي ليصبح مجموع العينة $(3 + 7 = 10)$ أشخاص. ومن العوامل التي تشجع على العينة الحصصية توفير الوقت والكلفة والجهد، والحصول على إجابات سريعة من العينة، كما أنها تصبح ضرورة عندما يكون شريحة في المجتمع ذات تمثيل قليل ونرغب في إشراكها في العينة المختارة.

4. عينة كرة الثلج Snowball Sample

تستخدم عينة كرة الثلج عندما نواجه صعوبة في تحديد أعضاء المجتمع المرغوب دراسته، حيث يبدأ الباحث بعينة صغيرة ميسرة، ثم تبدأ العينة بالكبر شيئاً فشيئاً مع سير الدراسة. وفي هذه الحالة نحتاج إلى الخطوات التالية: (Saunders et al., 2007, 232)

- الاتصال بواحد أو اثنين من حالات المجتمع المرغوب دراسته.
- سؤال هؤلاء لتحديد حالات أخرى يمكن الرجوع إليها لتوفر المعلومات لديهم.

- سؤال الحالات الجديدة لتحديد حالات أخرى جديدة، وهكذا.
- التوقف عندما لا نستطيع الوصول إلى حالات جديدة، أو الوصول إلى حجم عينة مقبول.

ومثال ذلك: إذا أراد باحث أن يدرس تأسيس الإمارة عام 1921 عن طريق المقابلات مع الأفراد الذين عايشوا الحدث، ونلاحظ في هذه الحالة أن الأفراد الذين عايشوا الحدث ولا زال عسى قيد الحياة قد يكون عددهم قليل، ولذلك يقوم الباحث بتحديد والاتصال بواحد أو اثنين من هؤلاء الأفراد، ثم يقوم بالاستدلال منهم على أفراد آخرين وهكذا حتى لا يستطيع الوصول إلى أفراد جدد، أو يكون قد استوفى البيانات التي يرغب بجمعها لبحثه.

ثانياً: العينات الاحتمالية Probabilistic Samples:

عينات يتم اختيارها بطرق تعتمد مبادئ الاحتمالات بغرض تمثيل المجتمع، ومنها ما يلي:

1. العينات العشوائية البسيطة Simple Random Samples:

اختيار عسدد معين من أفراد المجتمع بحيث يكون لأي فرد من الأفراد الفرصة نفسها للظهور في هذه العينة، وتستخدم للمجتمع الذي يتكون من عناصر متجانسة.

$$\text{حجم العينة} = \text{نسبة العينة} * \text{عدد أعضاء المجتمع}$$

2. العينات العشوائية الطبقة Stratified Random Sample:

يتم الحصول عليها بتقسيم المجتمع الأصلي إلى طبقات أو فئات وفقاً لخاصية معينة كالجنس أو مستوى التعليم، فإذا كانت عناصر المجتمع غير متجانسة فإننا نقسم المجتمع إلى طبقات، ثم نأخذ عينة عشوائية بسيطة من كل طبقة تتناسب مع حجم الطبقة. ويمكن تقسيم العينة الطبقة العشوائية إلى:

- توزيع متساوي (Equal Distribution) وهنا تقسم العينة الكلية على الطبقات بالتساوي.

• توزيع متناسب/ نسبي (Proportional Distribution) حيث يؤخذ عدداً من كل طبقة يناسب مع حجم الطبقة في المجتمع.

$$\text{العينة الطبقة} = (\text{حجم الطبقة} / \text{حجم المجتمع}) * \text{حجم العينة}$$

مثال: رأي الآباء والأمهات حول قضية معينة.

89-2003- أردنا اختيار عينة طبقية حجمها ن = 100 من مجتمع مكون من 1000 شخص وينقسم إلى طبقتين (400 ذكور، 600 إناث)، فإن هذه العينة ستكون مكونة من:

أ) 40 ذكور، 60 إناث ب) 60 ذكور، 40 إناث

ج) 50 ذكور، 50 إناث د) 30 ذكور، 70 إناث

العينة الطبقة = (حجم الطبقة / حجم المجتمع) * حجم العينة

حجم الذكور = $(1000/100) * 400 = 40$

حجم الإناث = $(1000/100) * 600 = 60$

85-2003- في دراسة إحصائية استهدفت طلبة كليات المجتمع، أخذت عينة عشوائية

من كل كلية يتناسب عددها مع عدد الطلبة فيها، فإن هذه العينة تسمى:

أ) عنقودية ب) منتظمة ج) معيارية د) طبقية.

81-2005- كلية تضم عدة تخصصات مختلفة، يراد اختيار عينة تمثل كل الطلاب في

الكلية، فإن أفضل أسلوب لاختيار هذه العينة هو العينة العشوائية:

أ) البسيطة ب) المنتظمة ج) الطبقة د) العنقودية.

2-2002- عدد المهندسين المسجلين في نقابة المهندسين 35000 مهندس

و 15000 مهندسة، وأردت اختيار عينة عددها (500) مهندس ومهندسة، فالطريقة الأنسب

في اختيار هذه العينة على أساس نقابي هي العينة:

أ- العشوائية. ب- المنتظمة. ج- العنقودية. د- الطبقة.

3. العينات العشوائية العنقودية Cluster Random Samples:

عندما نواجه في بعض الدراسات التطبيقية أن وحدات بعض المجتمعات تكون على شكل تجمعات وغالباً ما تكون متشابهة إلى حد كبير بالنسبة للخاصية التي نقوم بدراستها مثل: المدن، الشوارع، الكليات، وغيرها فإن هذه التجمعات عندها تسمى عناقيد (Cluster) إذ يحوي كل عنقود منها على عدد من عناصر المجتمع الأصلية والتي غالباً ما تكون متجانسة، فإننا نلجأ في هذه الحالة إلى العينة العنقودية.

تتميز مجموعات الدراسة المختلفة في المعاينة العنقودية بعدم التجانس بين عناصر كل مجموعة، حيث يوجد اختلافات بين العناصر المشكلة للمجموعة الواحدة، مع وجود تجانس بين المجموعات الجزئية (العناقيد). أي تجانس بين العناقيد ككل، ولكن عدم تجانس داخل العنقود نفسه.

وتقسم العينة العنقودية إلى:

- عينة عنقودية بمرحلة واحدة Single-Stage Cluster Sampling
- عينة عنقودية بمرحلتين Duple-Stage Cluster Sampling
- عينة عنقودية متعددة المراحل Multi-Stage Cluster Sampling
- معاينة مساحية Area Cluster Sampling

4. العينات العشوائية المنتظمة Systematic Random Samples:

وهي نادرة الاستعمال وتتصف بانتظام الفترات بين وحدات الاختيار، أي أن الفرق بين كل اختيار والدي يليه متساوياً في كل الحالات، ويستعمل إذا توفرت قائمة بأسماء أفراد المجتمع فإننا نستطيع اختيار أفراد العينة بحيث يكون الفرد ذو ترتيب معين ضمن أفراد المجتمع ويكون اختيار الفرد الأول من القائمة عشوائياً، مثال: اختيار (100) طالب من أصل (1000) طالب في الجامعة.

محدد مقدار الفترة = عدد طلاب المجتمع / عدد طلاب العينة = $10 = 100 / 1000$

تحديد نقطة البدء ويتم اختيارها عشوائياً من 0-9 ولنفترض الرقم 8 هو الاختيار

الأول، ونزيد لكل اختيار يليه الرقم 10

5. العينات المعيارية Standard Samples:

عينة تمثل المجتمع الإحصائي تمثيلاً صادقاً وتتفق مقاييسها الإحصائية مع مقاييس المجتمع (الوسط، الوسيط، الانحراف المعياري) ويتم اختيارها بصورة تنبؤية.

2004-22- من العينات الاحتمالية العشوائية:

أ) القصدية. ب) الحصصية. ج) العنقودية. د) الصدفة.

أخطاء العينات Samples Errors

إن النتائج التي نحصل عليها من العينات لا تكون مطابقة تماماً للنتائج في حالة المسح الشامل وذلك لأن نتائج العينات تتعرض لمجموعة من الأخطاء منها ما يلي:

1. أخطاء عشوائية (أخطاء الصدفة) Random Errors: والسبب في هذا الخطأ هو طريقة اختيار العينة، مثل اختيار (حجم العينة، نوع العينة، تباين عناصر المجتمع).

2. أخطاء التحيز Bias Errors: سببه زيادة أو نقص في البيانات، وقد يحدث هذا الخطأ أيضاً في المسح الشامل وذلك للأسباب التالية:

- أ- الإجابات الخاطئة التي قد يتسبب فيها جامع البيانات.
- ب- أخطاء من قبل المستجيب لعدم فهمه السؤال.
- ج- أخطاء من قبل المستجيب لأمر شخصية.
- د- التحيز في عناصر المجتمع التي تم اختيارها.
- هـ- عدم الوصول إلى مفردات العينة واستبدالها بأخرى.
- و- عدم وجود إطار سليم للعينة.

1-2002- يقصد بالعينة:

- أ- المشاهدات التي يتم تطبيقها على جميع أفراد مجتمع الدراسة.
- ب- مجموعة جزئية من مجتمع الدراسة.
- ج- إحدى وسائل المسح الشامل.
- د- طريقة إحصائية في قياس التزعة المركزية.

7-2002- يمكن الحكم على مدى تمثيل عينة ما للمجتمع المأخوذة منه من خلال:

أ- تجانس أفراد عينة الدراسة.

ب- تمثيل العينة بنسبة تزيد على 10%

ج- بعد أو قرب متوسط العينة عن متوسط مجتمعها مقدراً بوحدات الخطأ المعياري.

د- العينة منتظمة.

8-2002- أفضل نسبة في اختيار عينة الدراسة من مجتمع كبير برأي الإحصائيون:

أ- 2% ب- 4% ج- 5% د- 10%

72-2005- العينة الأكثر دقة في تمثيل المجتمع غير المتجانس هي:

أ- العشوائية البسيطة. ب- المنتظمة. ج- الطبقة. د- متعددة المراحل.

73-2005- مستوى القياس الذي تستخدم فيه الأرقام بهدف التصنيف فقط هو:

أ- الاسمي. ب- الرتبي. ج- النسبي. د- الفئوي.

92-2002- العينة التي تمنح كل فرد من أفراد المجتمع نفس الفرصة في الاختيار ليكون أحد

أفرادها هي العينة:

أ- الطبقة. ب- العشوائية البسيطة. ج- المنتظمة. د- متعددة المراحل.

93-2006- مستوى القياس الذي تكون وحداته متساوية وليس له صفر مطلق هو:

أ- النسبي. ب- الرتبي. ج- الفئوي. د- الاسمي.

100-2006- المقصود بمجتمع الدراسة في الإحصاء:

أ) الأفراد الذين تجرى عليهم الدراسة. ب) العينة التي تقع عليها الدراسة.

ج) الأفراد الذين تعميم عليهم نتائج الدراسة. د) جزء من عينة الدراسة.

8-2 توزيع المعاينة Sampling Distribution

يخضع المجتمع الذي تؤخذ منه العينة لتوزيع معين وهو توزيع المجتمع الاحتمالي لمتغير عشوائي يمثل وحدات ذلك المجتمع، وان التوزيع الاحتمالي للإحصاء يدعى بتوزيع المعاينة لتلك الإحصاءة والمتمثل بثوابت تعين هذا التوزيع تماماً وتسمى معلمات.

إذا كان المجتمع يخضع لتوزيع طبيعي فإن المعلمين هما الوسط الحسابي μ والانحراف المعياري σ فإذا كانت معلومة فيمكن عندئذ إيجاد جميع الاحتمالات المتوقعة لهذا المجتمع. وكذلك الحال إن كان المجتمع يخضع لتوزيع ذو حدين فإن المعلمة هي احتمال النجاح (p) ، فإذا علمت (p) فإننا نستطيع معرفة توزيعه أي يمكن تحديد مجتمعه، أي أن دالة الكثافة الاحتمالية للمتغير (x) والذي يمثل أي فرد من أفراد ذلك المجتمع يكون قد تحدد تماماً، وبعد ذلك يمكن حساب بعض المقاييس عن هذه العينة مثل الوسط الحسابي للعينة الواحدة هو (\bar{x}) ويسمى بإحصاءة العينة وهذه القيمة ربما تتغير من عينة إلى أخرى، وأن قيمة الوسط كما ذكرنا فإنها تتغير من عينة إلى أخرى لذا فالمتغير العشوائي هنا هو (\bar{x}) أي أن قيمة هذا الإحصاءة (\bar{x}) تتغير بتغير العينة، ولهذا يسمى التوزيع الاحتمالي لإحصاءة العينة، ومنها يمكن التوصل إلى تعريف الانحراف القياسي لتوزيع المعاينة والذي يدعى بالخطأ القياسي للإحصاءة.

نظرية: إذا كان (x) يخضع لتوزيع وسطه μ وتباينه σ^2 وكان (\bar{x}) يمثل الوسط الحسابي لعينة حجمها n والمسحوبة من هذا المجتمع فإن القيمة المتوقعة لهذا الوسط هو $\mu = \bar{x}$ أي أن الوسط الحسابي لـ (\bar{x}) هو الوسط الحسابي لجميع الأوساط الحسابية لعينات التي سحبت منها هذه العينات، أما تباين (\bar{x}) هو:

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$$

أي أن تباين هذه العينات يعتمد على تباين المجتمع وعلى حجم العينة، وهو بذلك أقل من تباين المجتمع، وبالتالي كلما كبر حجم العينة قل مقدار الخطأ القياسي للوسط الحسابي (\bar{x}) ونقرب وسط تلك العينة من الوسط الحسابي للمجتمع لذا يمكن استخدام تقدير الوسط الحسابي كتقدير لـ μ ، ويجب توفر شرط الإرجاع.

توزيع المعاينة للوسط الحسابي من مجتمع طبيعي:

Sampling Distribution for the Mean of Normal Population:

إذا سحبت عينة عشوائية حجمها (n) من مجتمع كبير له وسط حسابي (5) وتباين معلوم σ^2 فإن توزيع المعاينة للوسط الحسابي (\bar{x}) يقترب من التوزيع الطبيعي بوسط حسابي $\mu_{\bar{x}} = \mu$.

والانحراف المعياري:

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$$

وأن (Z) هي قيمة المتغير الطبيعي القياسي الذي له وسط حسابي مساوياً صفر وتباين مقداره واحد.

وهذا ما يسمى بتوزيع المعاينة للوسط الحسابي (\bar{x}).

مثال (2):

تخضع علامات الطلاب في مادة الإحصاء لتوزيع طبيعي بمعدل (70) والانحراف المعياري (20)، سحبت منه عينة عشوائية حجمها (36) طالباً، أوجد:

- توزيع المعاينة لهذه العينة.

- احسب احتمال أن يزيد معدل علامات الطلاب عن (78).

الحل:

$$\mu_{\bar{x}} = \mu = 70$$

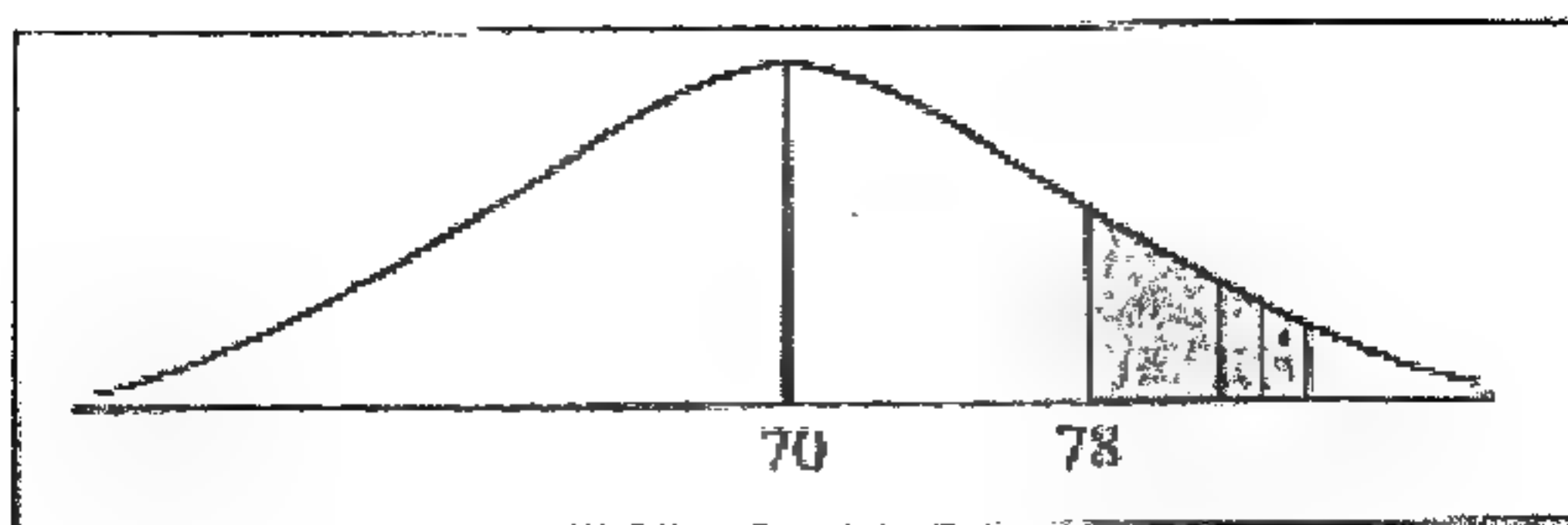
$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{20}{\sqrt{36}} = \frac{20}{6} = 3.3$$

وإن توزيع المعاينة هو:

$$\bar{X} \sim N(70, 10.89)$$

لإيجاد الاحتمال لدينا:

$$\begin{aligned} p(\bar{x} \geq 78) &= p\left(\frac{\bar{x} - 70}{3.3} \geq \frac{78 - 70}{3.3}\right) \\ &= p(Z \geq 2.42) \end{aligned}$$



$$P(Z \geq 2.42) = 0.5 - p(0 < Z < 2.42)$$

وباستخدام جدول التوزيع الطبيعي:

$$P(Z \geq 2.42) = 0.5 - 0.4922 = 0.0074$$

توزيع المعاينة للفرق بين وسطين:

Sampling Distribution for the Difference Between Two Sample Means:

في حالة معلومية تباين المجتمعين الموزعين توزيعاً طبيعياً

نفرض أن لدينا مجتمعين، الأول بوسط μ_1 وتباينه σ_1^2 والمجتمع الثاني وسطه الحسابي

μ_2 وتباينه σ_2^2 والمجتمعين يتبعان التوزيع الطبيعي:

نظرية: سحبت عينة عشوائية حجمها (n_1) من مجتمع طبيعي معدله μ_1 وتباينه σ_1^2 ،

وعينة ثانية من مجتمع طبيعي أيضاً معدله μ_2 وتباينه σ_2^2 والعينة الثانية مستقلة عن العينة

الأولى، فإذا كان (\bar{X}_1) يمثل الوسط الحسابي للعينة الأولى، و (\bar{X}_2) يمثل الوسط الحسابي للعينة

الثانية، فإن توزيع المعاينة للفرق بين متوسطيهما $(\bar{X}_1 - \bar{X}_2)$ يتبع التوزيع الطبيعي ذا المعدن

$(\mu_1 - \mu_2)$ والتباين:

$$\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

فإن توزيع المعاينة يكون:

$$(\bar{X}_1 - \bar{X}_2) \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)$$

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

يخضع لتوزيع طبيعي معياري بوسط مساوي إلى الصفر وانحراف معياري يساوي واحد، أي أن:

$$\mu_{\bar{X}_1 - \bar{X}_2} = \mu_1 - \mu_2$$

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

مثال (3):

سحبت عينيتين عشوائيتين من شركتين مختلفتين لإنتاج العدد الزراعي وكانت الأجر المدفوعة من قبل الشرطتين تتبع التوزيع الطبيعي، وإن معدل الأجر المدفوعة من قبل الشركة (A) إلى (36) عاملاً تساوي ديناراً أردنياً بانحراف معياري (36) ديناراً، أما الأجر المدفوعة من قبل الشركة (B) إلى (49) عاملاً تساوي (186) ديناراً بانحراف معياري (40) ديناراً، احسب احتمال أن معدل الأجر المدفوعة من قبل الشركة (A) لها متوسط على الأقل (60) ديناراً فأكثر من متوسط الأجر المدفوعة للشركة (B).

الحل:

المعلومات المتوفرة في المثال يمكن توضيحها كما يلي:

	الشركة (B)	الشركة (A)
الوسط الحسابي	$\mu_2 = 180$	$\mu_1 = 230$
الانحراف المعياري	$\sigma_2 = 40$	$\sigma_1 = 36$
حجم العينة	$n_2 = 49$	$n_1 = 36$

$$\begin{aligned}\mu_{X_1-X_2} &= \mu_1 - \mu_2 \\ &= 230 - 180 \\ &= 50\end{aligned}$$

$$\begin{aligned}\sigma_{X_1-X_2} &= \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \\ &= \sqrt{\frac{(36)^2}{36} + \frac{(40)^2}{49}} = 8.3\end{aligned}$$

$$P(\bar{X}_1 - \bar{X}_2 \geq 60) = P\left\{\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \geq \frac{60 - 50}{8.3}\right\}$$

$$\begin{aligned}P\left(Z \geq \frac{10}{8.3}\right) &= P(Z \geq 1.204) = 0.5 P(0 \leq Z \leq 1.204) \\ &= 0.5 - 0.3925 = 0.1075\end{aligned}$$

2- توزيع المعاينة للنسب Sampling Distribution for Proportion

إذا كانت قيمة كل عنصر متمثلة بالنجاح أو الفشل فإننا نسمي كل مشاهدة من هذا المجتمع "تجربة بيرنوللي" حيث أن احتمال النجاح يساوي (p) واحتمال الفشل يساوي (q) عملاً بأن (p + q = 1) وعندما نسحب عينة عشوائية حجمها (n) فإنه يجب إعادة التجربة (n) من المحاولات، وبذلك فإن توزيع المعاينة للمتغير العشوائي (X) المتمثل بعدد النجاحات في العينات ذات حجم (n) يمكن أن يكون قريباً من التوزيع الطبيعي بوسط حسابي قدره $\mu = np$ وانحراف قياسي $\sigma = \sqrt{npq}$ علماً بأن نسبة النجاح مختلفة من عينة إلى أخرى وعلى شرط ألا تكون قريبة من الصفر أو الواحد.

يمكن التعبير عن نسبة النجاح (p) بالمقدار: $\hat{P} = \frac{x}{n}$.

حيث أن (x) = عدد المحاولات (النجاحات).

n = حجم العينة.

\hat{p} = نسبة النجاح في العينة.

وكما ذكرنا أن \hat{p} تختلف من عينة إلى أخرى فإن توزيع المعاينة لـ (X) عدد النجاحات في العينات ذات الحجم (n) يمكن أن يكون قريباً من التوزيع الطبيعي بوسط حسابي قدره:

$$\mu_{\hat{p}} = E(\hat{p}) = E\left(\frac{X}{n}\right) = \frac{np}{n} = p$$

وتباين:

$$\sigma_{\hat{p}}^2 = \text{var}(\hat{p}) = \text{var}\left(\frac{X}{n}\right) = \frac{npq}{n^2} = \frac{pq}{n}$$

على شرط أن تكون قيم (p) قريبة من الصفر أو الواحد.

نظرية:

سحبت عينة عشوائية حجمها (n) من مجتمع إحصائي يخضع لتوزيع بيرنولي، أي ذات الحدين $B(n, p)$ بوسط $\mu = np$ وتباين $\sigma^2 = npq$ ، فتوزيع المعاينة إلى \hat{p} نسبة النجاحات هو قريب من التوزيع الطبيعي بوسط $\mu_{\hat{p}} = p$ وانحراف قياسي:

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$$

لذا فإن القيمة المعيارية (Z) هي:

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

$$\hat{p} = \frac{X}{n}$$

مثال (4): إذا كان احتمال نجاح الطالب في مادة الاحصاء هو 0.8 وسحبت عينة عشوائية حجمها 49 طالباً، أوجد احتمال $(0.7 < \hat{p} < 0.92)$

الحل: باستخدام النظرية يمكن إيجاد قيم الاحتمال أعلاه كما يلي:

$$P(0.7 \leq \hat{p} \leq 0.92)$$

$$P\left\{\frac{0.7 - 0.8}{\sqrt{\frac{(0.8) \times (0.2)}{49}}} \leq Z \leq \frac{0.92 - 0.8}{\sqrt{\frac{(0.8) \times (0.2)}{49}}}\right\}$$

$$= P\left\{\frac{-0.1}{0.06} \leq Z \leq \frac{0.12}{0.06}\right\}$$

$$= P(-1.66 \leq Z \leq 2) = P(0 \leq Z \leq 2) + P(0 < Z \leq 1.66) = 0.4772 + 0.4515 = 0.9287$$

توزيع المعاينة للفرق ما بين نسبتين:

Sampling Distribution for Differences Between Two Proportions:

نظرية: سحبت عيتان عشوائيتان حجمها (n_1, n_2) من مجتمعين مستقلين يخضع الأول

لتوزيع ذي الحدين $B(n_1, p_1)$ ، والثانية أيضاً تخضع لتوزيع ذو حدين $B(n_2, p_2)$ ، وأن:

$$\mu_1 = n_1 p_1, \quad \sigma_1^2 = n_1 p_1 q_1$$

$$\mu_2 = n_2 p_2, \quad \sigma_2^2 = n_2 p_2 q_2$$

فتوزيع المعاينة للفرق ما بين $\hat{P}_1 - \hat{P}_2$ يقترب من التوزيع الطبيعي بوسط:

$$\mu_{\hat{P}_1 - \hat{P}_2} = p_1 - p_2$$

وانحراف قياسي:

$$\sigma_{\hat{P}_1 - \hat{P}_2} = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

لذا فإن القيمة المعيارية لهما:

$$Z = \frac{\hat{P}_1 - \hat{P}_2 - (p_1 - p_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}}$$

مثال (5):

إذا كانت نسبة النجاح في مادة الإحصاء في جامعة مؤته تساوي (0.8)، وكانت نسبة النجاح لنفس المادة في جامعة الاسراء تساوي (0.75)، سحبت عينة عشوائية حجمها (70) طالباً من جامعة مؤته وعينة ثانية من جامعة الاسراء حجمها (35). أوجد احتمال أن تزيد نسبة النجاح في جامعة مؤته عن نسبة النجاح في جامعة الاسراء بمقدار (0.1) على الأكثر.

الحل:

$$\text{المطلوب هو إيجاد } P(\hat{P}_1 - \hat{P}_2 \leq 0.10)$$

وباستخدام النظرية أعلاه فإن هذا الاحتمال يساوي:

$$p(\hat{p}_1 - \hat{p}_2 \leq 0.10)$$

$$= p \left\{ \frac{\hat{P}_1 - \hat{P}_2 - (p_1 - p_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}} \leq \frac{0.10 - (0.8 - 0.75)}{\sqrt{\frac{(0.8) \times (0.2)}{70} + \frac{(0.75) \times (0.25)}{35}}} \right\}$$

$$= p \left(Z \leq \frac{0.05}{0.0873} \right) = p(Z \leq 0.573)$$

$$= 0.5 + P(0 \leq z \leq 0.573)$$

$$= 0.7157$$

أي أن نسبة النجاح في مؤته تزيد بـ (71%) عنها في جامعة الاسراء.

توزيع المعاينة للتباين (Sampling Distribution for the Variance):

إذا سحبنا عينات عشوائية كل منها ذات حجم n من مجتمع طبيعي تباينه σ^2 ثم أعيد الاختيار لعدة مرات وحسب تباين كل عينة S^2 فإننا نحصل على الإحصاءة S^2 .

فإن المعاينة S^2 ليس ذا مكانة عملية في الإحصاء لذا ننتج إلى توزيع المعاينة لـ:

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2}$$

هذا يعني أن توزيع المعاينة إلى التباين يخضع لتوزيع مربع كاي وبدرجة حرية $(n-1)$.

نظرية: مسحبت عينة عشوائية حجمها (n) من توزيع طبيعي معدله (μ) وتباينه أي $X \sim N(\mu, \sigma^2)$ وكان S^2 هو تباين العينة، لذا فإن:

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2}$$

هي قيمة المتغير العشوائي (χ^2) الذي له توزيع مربع كاي بدرجة حرية $(n-1)$.

مثال: تخضع علامات الطلاب في مادة القياس والتقويم لتوزيع طبيعي بمعدل 70

والمحراف المعياري 10 مسحبت منه عينة عشوائية حجمها 25 طالباً.

- جد توزيع المعاينة لهذه العينة.

- احسب احتمال أن يزيد معدل علامات الطلاب عن 78

$$\mu_{\bar{x}} = \mu = 70$$

$$\sigma_{\bar{x}} = \sigma / \sqrt{n} = 10 / \sqrt{25} = 10/5 = 2$$

ان توزيع المعاينة هو:

$$\bar{x} \sim N(\mu, \sigma^2/n)$$

$$\bar{x} \sim N(70, 100/25), \bar{x} \sim N(70, 4)$$

احتمال أن يزيد معدل علامات الطلاب عن 78

$$P(\bar{x} \geq 78) = P\left\{\frac{\bar{x} - 70}{2} \geq \frac{78 - 70}{2}\right\} \\ = P(Z \geq 4)$$

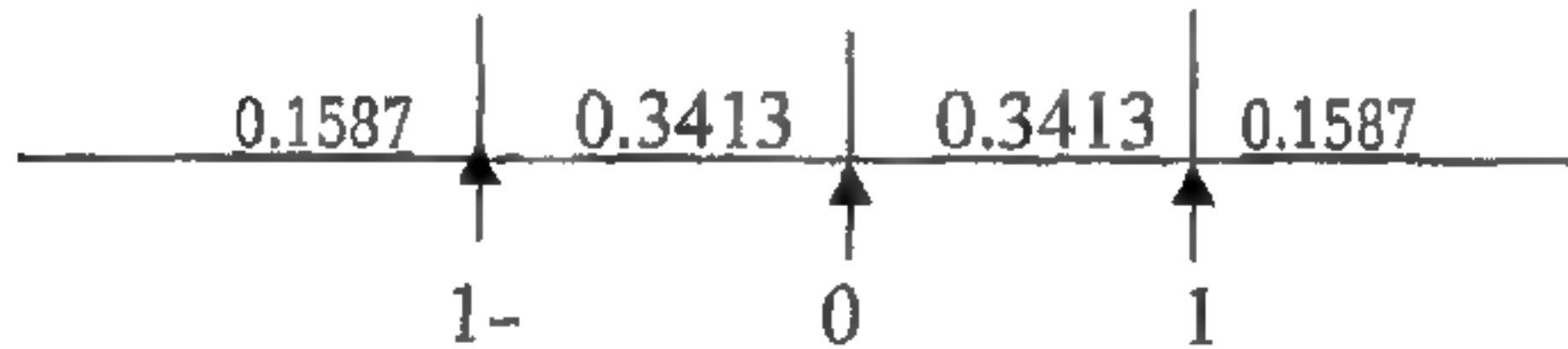
9-2 أسئلة وتمارين Exercise

س1: ضع رمز الإجابة الصحيحة في المربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										

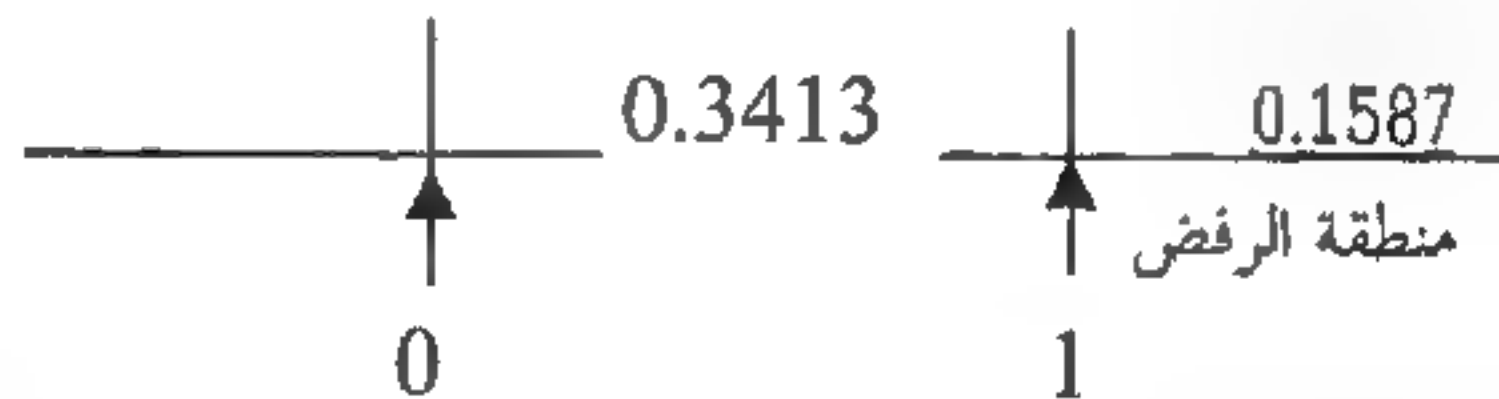
1- في المنحنى الطبيعي المعياري المساحة المحصورة بين $Z=+1$ و $Z=-1$ هي:

أ- 0.1587 ب- 0.3413 ج- 0.6826 د- 0.3174



2- إذا كان $X=60$ ، $S_x=10$ ، حيث X علامة الإحصاء، ما احتمال الحصول على علامة أكبر من 70 ؟

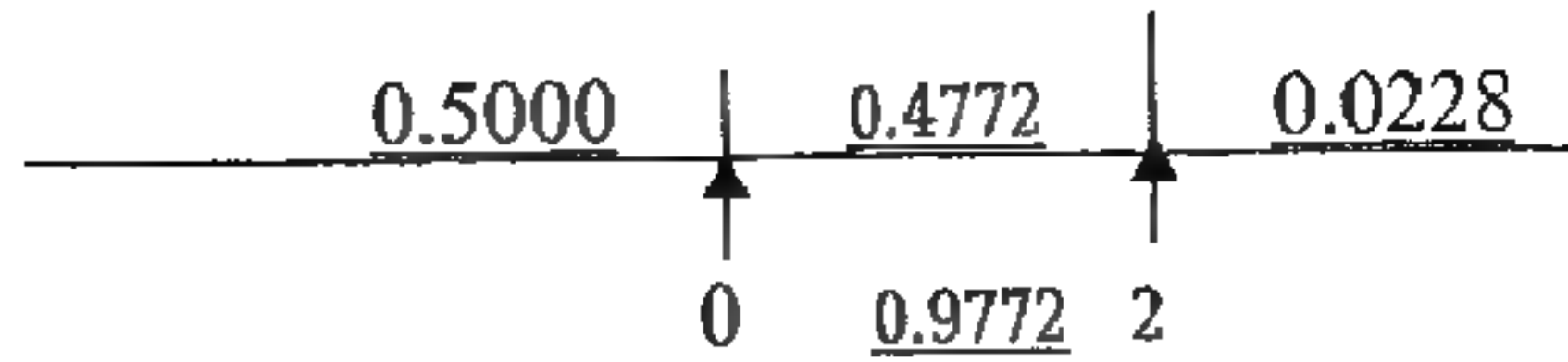
أ- 0.3413 ب- 0.5000 ج- 0.1587 د- 0.8413



3- تضم جامعة (10000) طالباً وطالبة وتوزع أوزانهم توزيعاً طبيعياً معيارياً بوسط حسابي

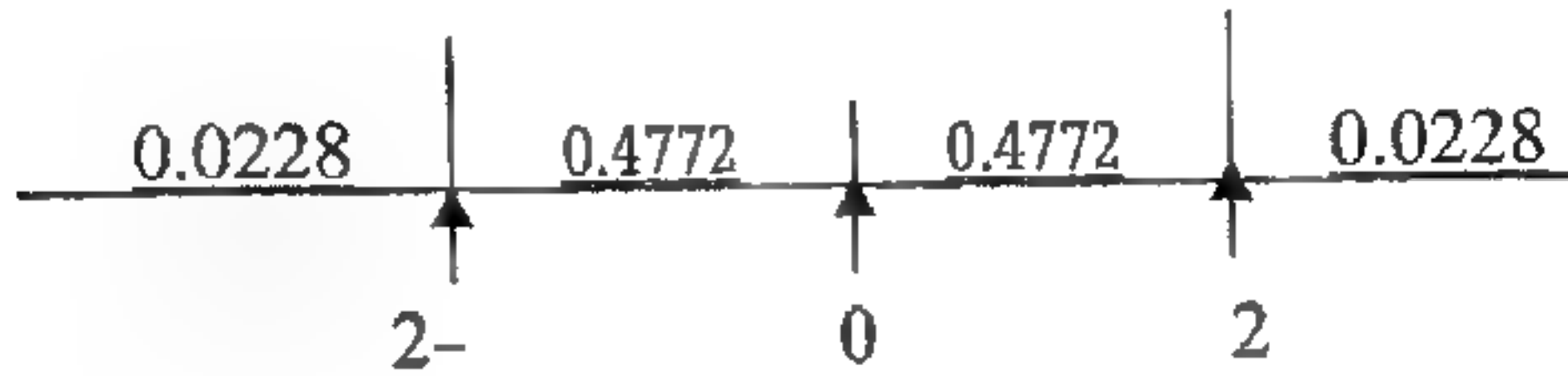
$X = 60$ والانحراف المعياري $S_x = 10$ ، ما عدد الطلبة الذين تقل أوزانهم عن 80 كغم؟

أ- 0.9772 ب- 0.5000 ج- 0.0228 د- 0.4772



4- استعن بالجدول لإيجاد X' إذا علمت أن $X = 16$ ، $S_x = 5$:

- أ- 6 ب- 2 ج- (6-) د- (2-)



5- إذا تم تحويل علامات من مقياس إلى آخر ولم يصاحب التحويل تغييراً في شكل التوزيع فإن التحويل من النوع.....

6- العلامة في توزيع وسطه الحسابي = 0، وانحرافه المعياري = 1 تسمى العلامة المعيارية

7- العلامة في توزيع وسطه الحسابي = 50، وانحرافه المعياري = 10 تسمى العلامة المعيارية

* إذا كان لتوزيع وسط حسابي = 30، وانحراف معياري = 5 وكانت لعلامة طالب انحرافاً معيارياً واحداً دون الوسط الحسابي فإن:

8- العلامة الخام للطالب =

9- العلامة الناتجة للطالب =

10- يسعى الباحث إلى تحويل العلامات من توزيع إلى آخر بغرض

س2: إذا كانت درجات الطلاب في امتحان الرياضيات تتبع التوزيع الطبيعي، وان متوسط الدرجات هو 60 وانحرافها المعياري 10 درجات، ما هو احتمال أن تكون درجة احد الطلاب : الدرجة تقع بين $60 \leq X < 70$ ، الدرجة تقع بين $50 < X < 85$

س3: إذا كان احتمال رسوب الطالب في مادة الإحصاء هو 0.40 وأخذت عينة حجمها $n =$

32 طالب من أولئك الذين يدرسون مادة الإحصاء، فما هو احتمال $P \leq 0.30$

س4: إذا كانت نسبة الإناث في جامعة اليرموك هي 0.40 ونسبتهم في جامعة مؤتة هي

0.60 وأخذت عينة من طلبة جامعة اليرموك حجمها 100 وعينة من طلبة جامعة

مؤتة حجمها 200 فما هي صيغة التوزيع التقريبي للفرق بين نسبي العيتين مع إيجاد

احتمال $P[|p_2 - p_1| \leq 0.1]$

س5: إذا كانت الاحصاء اليومية لعمال إحدى الشركات تخضع لتوزيع طبيعي $N(5, 0.5)$

وكانت الاحصاء اليومية لعمال شركة أخرى مماثلة في النشاط تخضع لتوزيع طبيعي

$N(5, 0.25)$ ، وأخذت عينة من الشركة الأولى حجمها $n_1 = 20$ ومن الشركة الثانية

حجمها $n_2 = 10$ أوجد $P[x_1 - x_2 \leq 0.75]$

الفصل الثالث

التقدير وفترات الثقة

تقدير معالم المجتمع من معالم العينة

(باستخدام عينة واحدة)

Estimation of Population Parameters

1-3 مقدمة Introduction

2-3 خواص جودة التقدير

Properties of Goodness of Estimation

3-3 أنواع القيم التقديرية لمعالم المجتمع.

4-3 فترات الثقة Confidence Interval

5-3 العلاقة بين معلمات المجتمع وإحصائيات العينات.

6-3 تمارين Exercise

الفصل الثالث

التقدير وفترات الثقة

تقدير معالم المجتمع من معالم العينة (باستخدام عينة واحدة)

Estimation of Population Parameters

1-3 مقدمة Introduction

المجتمع Population: مجتمع بيانات أو مشاهدات أو علامات يحدد هويته الباحث، ويمثل جميع عناصر الظاهرة في الدراسة.

العينة Sample: مجموعة جزئية من المجتمع، أو من عناصر الظاهرة.

المؤشر Index: تدل على جميع مقاييس الترة المركزية والتشتت والعلاقة (الارتباط) سواء محسوبة لعينات أو لمجتمعات.

مؤشر عينة (إحصائي) Statistic: قيمة رقمية تصف خاصية معينة تعود للعينة مثل الوسط الحسابي لعينة \bar{X} .

مؤشر مجتمع (معلم) Parameter: قيمة رقمية تصف خاصية معينة تعود للمجتمع الإحصائي وتخص المجتمع، مثل الوسط الحسابي للمجتمع μ .

إن الهدف الأساسي من استخدام العينات هو إيجاد تقديرات تكون ممثلة لمعالم المجتمع Population Parameters عموماً مثل الوسط الحسابي والتباين والنسبة ومعامل الارتباط، وذلك لإمكانية استخدام نظرية الاحتمالات التي تقوم على أساسها هذه العينات العشوائية لدراسة الأخطاء المعيارية للتقديرات وإيجاد فترة الثقة لقيم معالم المجتمع الحقيقية بالاستناد على نتائج العينة، وعن هذه الطريقة نتمكن من تكوين فكرة عن مدى دقة ومقبولية هذه التقديرات.

العينات Samples

العينة Sample: أي مجموعة جزئية من المجتمع الإحصائي يتم جمع البيانات من خلالها بصورة مباشرة.

2-3 خواص جودة التقدير

Properties of Goodness of Estimation

تتلخص معايير دقة التقدير بما يلي:

1- **عدم التحيز Unbiasness**: وهذا يشير إلى أن تقدير الوسط الحسابي \bar{X} لعينة يعود إلى مستغير عشوائي من حيث الحجم وعملية السحب من المجتمع مستوفياً لشروط العشوائية، أي أن: $E(\bar{X}) = \mu$

حيث أن: μ : متوسط المجتمع الذي سحبت منه العينة. \bar{X} : متوسط العينة.
فعند تساوي كلا الوسطين يكون التقدير غير متحيز.

2- **التناسق Consistency**: وهو قرب قيمة التقدير إلى قيمة المعلمة.

كلما ازداد حجم العينة يقال أن التقدير قد حقق التناسق مع العينة.

3- **الكفاءة Efficiency**: وتشير إلى التباين في العينة المستخدمة لأي تقدير، إن التقدير الأقل تباين يكون نسبياً الأكثر كفاءة.

3-3 أنواع القيم التقديرية لمعالم المجتمع.

في الغالب نعلم إلى اختيار عينة Sample من المجتمع Population وبعد ذلك نحاول دراسة هذه الظاهرة بالنسبة لأفراد تلك العينة، وبالتالي نعمل على تعميم ما نتوصل إليه من نتائج على كافة أفراد ذلك المجتمع.
تنقسم التقديرات إلى نوعين:

1- التقدير بنقطة Point Estimation

يتم التقدير هنا بنقطة واحدة أو قيمة واحدة محددة، حيث نقدر معلمة المجتمع بنقطة تحسب من بيانات العينة، مثلاً قيمة الوسط الحسابي \bar{X} باعتبارها تقديراً لمعلمة الوسط الحسابي للمجتمع μ ، إن تباين العينة S^2 لتقدير تباين المجتمع σ^2 ، حيث أن معلمة التباين من العينة هو عبارة عن تباين متوسط العينة كما يلي:

$$\sigma_{\bar{X}}^2 = \sigma^2 / n$$

$$S_{\bar{X}} = S / \sqrt{n} \quad \text{وأن الخطأ المعياري يقدر بالاحصاءة كما يلي:}$$

مثال: إذا كان لدينا عينة حجمها $n=5$ وقيم وحداتها هي 3,4,5,1,2 مسحوبة من مجتمع مجهول المعالم.

فإن الوسط الحسابي للعينة $\bar{x}=3$ هي نقطة التقدير لقيمة متوسط المجتمع μ غير المعنوم. وإذا افترضنا بأن نسبة عدد القيم الزوجية للعينة p هي المقدّر لنسبة المجتمع π ، فإن $p=2/5=0.4$ هي نقطة التقدير، أي قيمة المقدّر لنسبة المجتمع p ونفس القول عند استخدام S ليدل على التقدير للانحراف المعياري للمجتمع غير المعنوم σ .

2- التقدير بفتره Interval Estimation

إننا لا نتوقع الحصول على المعالم المقدرة بدون أخطاء مهما كانت دقة التقدير، مع أن التقدير يزداد ثقة بزيادة حجم العينة، لذا يجب إعطاء فترة معينة لتوقع وقوع معالم المجتمع داخلها، والمقصود هنا الفترة التي تشمل قيمة المعلمة المجهولة باحتمال معلوم، وهي ما يسمى بفترة الثقة، ويمكن تقدير μ بفترة لها مقدار ثقة معلوم.

3-4 فترات الثقة Confidence Interval

العلاقة بين معالم المجتمع وإحصائيات العينات. تحدد كما يلي:

1- تقدير المتوسط الحسابي μ .

قيمة متوسط المجتمع (μ) = المتوسط العام لمتوسطات تلك العينات أي أن:

$$\mu = (\bar{x}_1 + \bar{x}_2 + \bar{x}_3 + \dots + \bar{x}_k) / k$$

ك: عدد العينات المأخوذة من حجم معين من المجتمع. k

م: متوسط المجتمع. μ م1: متوسط العينة الأولى. \bar{x}_1

م2: متوسط العينة الثانية. \bar{x}_2 م ك: متوسط العينة الأخيرة. \bar{x}_k

* الأساليب الإحصائية المستخدمة في استخراج معلومات المجتمع اعتماداً على إحصائيات العينة:

1- المتوسط الحسابي Mean

لتقدير قيمة متوسط المجتمع (μ) من خلال متوسط عينة مأخوذة منه X' نلجأ إلى الأساليب التالية:

أ- التقدير من خلال قيمة واحدة.

ب- التقدير من خلال مجموعة من القيم.

أ- التقدير من خلال قيمة واحدة.

نعتبر قيمة متوسط المجتمع μ مساوية لقيمة متوسط العينة X' أي أن:

$$\mu = \bar{X}$$

ب- التقدير من خلال مجموعة من القيم.

$$\mu = \bar{X} \pm Z * SE$$

حيث:

μ : متوسط المجتمع.

\bar{X} : متوسط العينة.

SE: الخطأ المعياري.

Z : العلامة المعيارية المناظرة لمستوى الثقة المستخدم في عملية التقدير.

مثال: إذا كان متوسط أوزان عينة من الأفراد = 70 كغم، وان الخطأ المعياري للوزن =

5 كغم، وان مستوى الثقة المستخدم في عملية التقدير 95%، جد متوسط المجتمع.

$$\mu = \bar{X} \pm Z * SE = 70 \pm 1.96 * 5$$

والمعلومات الخاصة بالمجتمع غير معروفة في كثير من الأحيان، فإن قيمة الخطأ المعياري

يسندر أن تكون معروفة بشكل مسبق، في هذه الحالة يستعاض عنه بالانحراف المعياري للعينة،

وتصبح المعادلة التي يستخرج منها متوسط المجتمع كالتالي:

$$\mu = \bar{X} \pm T * (\sigma/\sqrt{n})$$

حيث:

μ : متوسط المجتمع. \bar{X} : متوسط العينة.

T : العلامة المعيارية المعدلة. وتستخرج من جداول التوزيع التائي الخاصة بها.

n : عدد أفراد العينة.

نحتاج لمعرفة T إلى معرفة: مستوى الثقة، حجم العينة، درجات الحرية = n-1

مستوى الثقة	حجم العينة	درجات الحرية	T
%95	30	29	2.462
%95	25	24	2.492
%95	20	19	2.539

مثال: عينة تتألف من 20 طالب من طلبة الصف الثالث وكان متوسط علاماتهم على اختبار في الرياضيات 80 بانحراف معياري مقداره 5 ، على اعتبار أن مستوى الثقة هو 95 %، جد متوسط أداء طلبة الصف الثالث على هذا الاختبار؟

$$\mu = \bar{X} \pm T * (\sigma/\sqrt{n}) = 80 \pm 2.539 * (5/\sqrt{20}) = 80 \pm 2.539 * (5/4.47) \\ = 80 \pm 2.84 = (82.84) , (77.16)$$

إذا فترة الثقة هي [77.16 – 82.84]

وبما أن قيم متوسطات العينات مختلفة فيما بينها حيث أن بعضها = متوسط المجتمع والآخر مختلف عنه، فإن درجة تشتت قيم هذه المتوسطات حول متوسط المجتمع والتي تسمى الخطأ المعياري وقانونه هو:

$$SE = \sigma / \sqrt{n}$$

وكلما زاد عدد أفراد العينة كلما قل الخطأ المعياري، وكلما اقتربت قيم متوسطات العينات من قيمة متوسط المجتمع، كلما قل الخطأ المعياري.

1- فترة الثقة للمتوسط μ Confidence Interval for μ أ- فترة الثقة لمتوسط مجتمع طبيعي μ تباينه σ^2 معلوم.

Confidence Interval for Population Means (Variance Known)

اختيار عينة عشوائية من مجتمع طبيعي، انحرافه المعياري σ معلوم، فإن متوسط العينة \bar{X} يتبع التوزيع الطبيعي بخطأ معياري σ/\sqrt{n} ، حيث أن n هي حجم العينة. وللمعرفة $(Z_{(1-\alpha/2)})$ نحتاج الى معرفة مستوى الثقة:

مستوى الثقة	قيمة Z
%95	1.96
%99	2.58

وعند ثقة %95 فإن المساحة التي يقع ضمنها المتوسط الحقيقي هي:

$$1.96 * (\sigma/\sqrt{n}) < \mu < 1.96 * (\sigma/\sqrt{n})$$

وان المتغير الطبيعي المعياري لتوزيع متوسط العينة Z هو :

$$Z = (\bar{X} - \mu) / (\sigma_{\bar{X}})$$

مثال: عند ثقة %95 فإن المنطقة التي يقع ضمنها المتوسط الحقيقي μ تقل بمقدار

$$1.96 * (\sigma/\sqrt{n})$$
 عن المتوسط الحقيقي وتزداد بمقدار $1.96 * (\sigma/\sqrt{n})$

وان المتغير الطبيعي المعياري لتوزيع متوسط العينة Z هو :

$$Z = (\bar{X} - \mu) / (\sigma_{\bar{X}}) \Rightarrow Z = (\bar{X} - \mu) / (\sigma/\sqrt{n})$$

حيث أن:

لم : متوسط المجتمع الذي سحبت منه العينة.

 \bar{X} : متوسط العينة. $\sigma_{\bar{X}}$: الانحراف المعياري للمتوسط ، أي σ/\sqrt{n}

إذا كانت درجة الثقة 100% فإن فترة الثقة تحسب من المعادلة التالية:

$$\bar{X} - (\sigma/\sqrt{n})(Z_{(1-\alpha/2)}) \leq \mu \leq \bar{X} + (\sigma/\sqrt{n})(Z_{(1-\alpha/2)})$$

مثال: مجتمع موزع توزيعاً طبيعياً انحرافه المعياري $\sigma = 11$ سحبت منه عينة عشوائية حجمها $n=26$ وكان متوسط العينة $\bar{x}=48$ احسب تقدير متوسط المجتمع μ بثقة 95%
الحل:

نستخرج قيمة الانحراف المعياري للمتوسط

$$\sigma_{\bar{x}} = \sigma/\sqrt{n} = 11/\sqrt{26} = 2.157$$

نستخرج القيمة الجدولية $Z_{(1-\alpha/2)}$

$$Z_{(1-\alpha/2)} = Z_{(1-0.05/2)} = Z_{(1-0.025)} = Z_{(0.975)} = 1.96$$

باستخدام القانون: $\bar{X} - (Z_{(1-\alpha/2)}) (\sigma/\sqrt{n}) \leq \mu \leq \bar{X} + (Z_{(1-\alpha/2)}) (\sigma/\sqrt{n})$

$$48 - (1.96) (2.157) \leq \mu \leq 48 + (1.96) (2.157)$$

$$43.77 \leq \mu \leq 52.23$$

إن تقدير متوسط المجتمع μ بثقة 95% هي في الفترة [43.77 - 52.23]

ب- فترة الثقة لمتوسط مجتمع طبيعي μ تبينه σ^2 غير معلوم وحجم العينة كبير.

Confidence Interval for Population Means (Variance unknown, Large Sample)

$$\bar{X} - (S/\sqrt{n}) (Z_{(1-\alpha/2)}) \leq \mu \leq \bar{X} + (S/\sqrt{n}) (Z_{(1-\alpha/2)})$$

مثال: سحبت عينة حجمها 64 طالب والمتمثلة بأعمار طلاب الجامعة وكان معدل العمر 20 عاماً والانحراف المعياري للعمر 10، قدر العمر الفعلي μ للمجتمع الطلاب بحدود ثقة 95%.

باستخدام القانون: $\bar{X} - (Z_{(1-\alpha/2)}) (S/\sqrt{n}) \leq \mu \leq \bar{X} + (Z_{(1-\alpha/2)}) (S/\sqrt{n})$

نستخرج القيمة الجدولية $Z_{(1-\alpha/2)}$

$$Z_{(1-\alpha/2)} = Z_{(1-0.05/2)} = Z_{(1-0.025)} = Z_{(0.975)} = 1.96$$

$$20 - 1.96 * (10/\sqrt{64}) \leq \mu \leq 20 + 1.96 * (10/\sqrt{64})$$

$$20 - 1.96 * 1.25 \leq \mu \leq 20 + 1.96 * 1.25$$

$$20 - 2.45 \leq \mu \leq 20 + 2.45 \Rightarrow 17.55 \leq \mu \leq 22.45$$

إن تقدير متوسط المجتمع μ بثقة 95% هي في الفترة [17.55 - 22.45]

ج- فترة الثقة لموسط مجتمع طبيعي μ تباينه σ^2 غير معلوم وحجم العينة صغير.

Confidence Interval for Population Means (Variance unknown, Small Sample)

$$S_{x'} = S/\sqrt{n} \quad \text{الخطأ المعياري التقديري لموسط العينة}$$

$$t = (\bar{X} - \mu) / (S_{x'}) = (\bar{X} - \mu) / (S/\sqrt{n})$$

$$\bar{X} - t_{(1-\alpha/2),v} * (S/\sqrt{n}) \leq \mu \leq \bar{X} + t_{(1-\alpha/2),v} * (S/\sqrt{n})$$

د- فترة الثقة لموسط مجتمع طبيعي μ تباينه σ^2 غير معلوم وتوزيعه غير معلوم.

$$S_{x'} = S/\sqrt{n} \quad \text{الخطأ المعياري التقديري لموسط العينة}$$

$$Z = (X' - \mu) / (\sigma_{x'})$$

$$\bar{X} - Z_{(1-\alpha/2)} * (S/\sqrt{n}) \leq \mu \leq \bar{X} + Z_{(1-\alpha/2)} * (S/\sqrt{n})$$

في حالة العينات التي تؤخذ من مجتمع والتي لها جميعها نفس عدد الأفراد فإن العلاقة بين معلومات المجتمع وإحصائيات العينات يمكن أن تحدد كما يلي:

2- قيمة تباين المجتمع (σ^2) = المتوسط العام لتباينات تلك العينات أي أن:

$$\sigma^2 = (\sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \dots + \sigma_k^2) / k$$

ويكون صحيحاً فقط إذا حسب تباين المجتمع من القانون:

$$\sigma^2 = \sum (X - \mu)^2 / N$$

N: عدد أفراد المجتمع.

وحسب تباين العينة من القانون:

$$\sigma^2 = \sum (X - X')^2 / n-1$$

n: عدد أفراد العينة.

التباين Variance:

لتقدير قيمة تباين المجتمع σ^2 من خلال تباين عينة مأخوذة منه S^2 نلجأ إلى الأساليب

التالية:

أ- التقدير من خلال قيمة واحدة.

ب- التقدير من خلال مجموعة من القيم.

أ- التقدير من خلال قيمة واحدة.

نعتبر قيمة تباين المجتمع σ^2 مساوية لقيمة تباين العينة S^2 أي أن:

$$S^2 = \sigma^2$$

ب- التقدير من خلال مجموعة من القيم. وتعتمد على:

مستوى الثقة = 1 - مستوى الدلالة = $1 - \alpha$

حجم مستوى الثقة المطلوب، قيمة الإحصائي (χ^2 / v)

وتقدر قيمة تباين المجتمع من خلال العلاقة التالية:

$$\frac{S^2}{\chi^2_{\frac{1-\alpha}{2}}(v)} \leq \sigma^2 \leq \frac{S^2}{\chi^2_{\frac{\alpha}{2}}(v)}$$

مثال: أخذت عينة عدد أفرادها 25 من بين الإناث في المرحلة الجامعية ووجد أن تباين أطوالهن هو 25 ، فمل هو تباين المجتمع في مثل هذه الحالة إذا كان مستوى الثقة المطلوب هو 99%

الحل:

$$S^2 = 25$$

$$V = 25 - 1 = 24$$

$$\alpha = 1 - 0.99 = 0.01$$

$$\chi^2_{\frac{\alpha}{2}}(v) = 0.412$$

$$\chi^2_{\frac{1-\alpha}{2}}(v) = 1.9$$

$$\frac{25}{1.9} \leq \sigma^2 \leq \frac{25}{0.412}$$

$$13.16 \leq \sigma^2 \leq 60.7$$

3- تقدير فترات الثقة للنسب في مجتمع Interval Estimation of Proportion

في حالة المجتمع: يرمز لنسبة وقوع الظاهرة في المجتمع ب P وهي نسبة النجاح.

ويرمز لنسبة عدم وقوع الظاهرة في المجتمع ب Q وهي نسبة الفشل، وهي (1-P).

أما في حالة العينة: يرمز لنسبة وقوع الظاهرة في العينة بـ p' وهي نسبة النجاح.
ويرمز لنسبة عدم وقوع الظاهرة في العينة بـ P وهي نسبة الفشل.
لتقدير قيمة نسبة المجتمع P من خلال نسبة عينة مأخوذة p منه نلجأ إلى الأساليب التالية:

أ- التقدير من خلال قيمة واحدة.

ب- التقدير من خلال مجموعة من القيم.

أ- التقدير من خلال قيمة واحدة.

نعتبر قيمة وجود النسبة لظاهرة ما في المجتمع P مساوية لقيمة وجودها في العينة المأخوذة من المجتمع p أي أن:

$$p = P$$

ب- التقدير من خلال مجموعة من القيم.

تقدير قيمة النسبة في المجتمع ضمن مجموعة من القيم.

$$P = (k/n) \pm Z * \sqrt{\frac{(k/n) * [(1-(k/n))]}{n}}$$

$$\hat{P} - (\sqrt{pq/n}) (Z_{(1-\alpha/2)}) < P < \hat{P} + (\sqrt{pq/n}) (Z_{(1-\alpha/2)})$$

P : النسبة في المجتمع. k/n : النسبة في العينة.

n : عدد أفراد العينة. k : عدد الحالات التي تنطبق عليها الظاهرة المدروسة.

Z : تعتمد على قيمة مستوى الثقة المطلوب.

مثال: أخذت عينة تتكون من 100 طالب جامعي من الذكور وسئلوا عن رأيهم في حرية الاختلاط في الحرم الجامعي، فأجاب 70 منهم بالإيجاب، فما هي النسبة العامة لتأييد الذكور لحرية الاختلاط في الحرم الجامعي، إذا كان مستوى الثقة المطلوب هو 99%.

الحل:

$$Z: 2.576$$

$$k: 70$$

$$n: 100$$

$$0.70 = 100/70 = k/n$$

$$P = (k/n) \pm Z * \sqrt{\frac{(k/n) * [(1-(k/n))]}{n}} = 0.70 \pm 2.576 * \sqrt{\frac{(0.70) * [0.3]}{100}}$$

$$= 0.70 \pm 2.576 * 0.046 = 0.70 \pm 0.12 = 0.82 , 0.58$$

4- معامل الارتباط Correlation Coefficient

ما المقصود بعبارة "إن معامل الارتباط بين متغيرين كان ذا دلالة إحصائية؟"

إن هذه العبارة استدلال إحصائي ينتج عن رفض الفرضية الصفرية $H_0: \rho = 0$ وهي تعني أن معامل الارتباط في المجتمع بين المتغيرين لا يختلف عن صفر، ورفض هذه الفرضية الصفرية يعني أن قيمة معامل الارتباط في المجتمع تختلف عن الصفر اختلافاً جوهرياً، أي أنها لم تنشأ عن خطأ المعاينة.

يستخدم توزيع المعاينة لمعامل الارتباط بيرسون r في اختبار الفرضية الصفرية $H_0: \rho = 0$ لتقدير قيمة نسبة المجتمع P من خلال نسبة عينة مأخوذة p منه نلجأ إلى الأساليب

التالية:

أ- التقدير من خلال قيمة واحدة.

ب- التقدير من خلال مجموعة من القيم.

أ- التقدير من خلال قيمة واحدة.

نعتبر قيمة معامل الارتباط لظاهرة ما في المجتمع r مساوية لقيمة وجودها في العينة المأخوذة من المجتمع r_1 أي أن:

$$r_1 r =$$

ب- التقدير من خلال مجموعة من القيم.

تقدير قيمة معامل الارتباط في المجتمع ضمن مجموعة من القيم.

$$Z(r) = Z(r_1) \pm Z * (1/\sqrt{n-3})$$

حيث:

r : معامل ارتباط المجتمع .

r_1 : معامل ارتباط العينة.

n : عدد أفراد العينة.

Z : القيمة المعيارية المناظرة لمستوى الثقة المطلوب والمستخرجة من الجدول.

$Z(r)$: القيمة الانحرافية لمعامل الارتباط (r) والمستخرجة من الجدول.

مثال: إذا كان معامل الارتباط بين الذكاء والتحصيل عند طلبة الجامعة هو 0.62 في حالة عينة أفرادها 67 طالباً وطالبة، فإن قيمة معامل الارتباط بين الذكاء والتحصيل في هذه الحالة الدراسية بوجه عام إذا كان مستوى الثقة المطلوب هو 0.95

الحل:

r : معامل ارتباط المجتمع .

$r: 0.62$

$n: 67$

$Z: 1.96$

$Z(r_1) = Z(0.62) = 0.73$ من الجدول

$$Z(r) = Z(r_1) \pm Z * (1/\sqrt{n-3}) = 0.73 \pm 1.96 * (1/\sqrt{67-3}) = 0.73 \pm 1.96 * 0.125 \\ = 0.73 \pm 0.245 = 0.98 , 0.48$$

من الجدول نجد أن r التي تناظر $Z(r) = 0.98$ هي 0.7531 أو 0.75

من الجدول نجد أن r التي تناظر $Z(r) = 0.48$ هي 0.4462 أو 0.45 وبذلك تكون

قيم الارتباط المطلوب هي: $r=0.75$ كحد أعلى ، $r=0.45$ كحد أدنى.

جدول (40): قيم Z الجدولية الموزعة توزيعاً طبيعياً $N(0,1)$ عند مستويات معنوية مختلفة.

P	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.50	0.000	0.025	0.050	0.075	0.100	0.126	0.151	0.176	0.202	0.228
0.60	0.253	0.279	0.305	0.332	0.358	0.385	0.412	0.44	0.468	0.496
0.70	0.524	0.553	0.583	0.613	0.643	0.674	0.706	0.739	0.772	0.806
0.80	0.842	0.878	0.915	0.954	0.994	1.036	1.080	1.126	1.175	1.227
0.90	1.282	1.341	1.405	1.476	1.555					
P	0.000	0.001	0.002	0.003	0.004	0.005	0.006	0.007	0.008	0.009
0.95	1.645	1.655	1.665	1.675	1.685	1.695	1.706	1.717	1.728	1.739
0.96	1.751	1.762	1.774	1.787	1.799	1.812	1.825	1.838	1.852	1.866
0.97	1.881	1.896	1.911	1.927	1.943	1.960	1.977	1.995	2.014	2.034
0.98	2.054	2.075	2.097	2.120	2.144	2.170	2.197	2.226	2.257	2.290
0.99	2.326	2.366	2.409	2.457	2.512	2.576	2.652	2.748	2.878	3.090

5-3 تمارين Exercise

س(1): ضع رمز الإجابة الصحيحة في المربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										

- 1- أي من التالية ليست من معايير دقة التقدير:
 - أ- عدم التحيز.
 - ب- التناسق.
 - ج- الموضوعية.
 - د- الكفاءة.
- 2- إن التناسق كمعيار من معايير دقة التقدير يعني:
 - أ- التغاير في العينة المستخدمة لأي تقدير.
 - ب- قرب قيمة التقدير إلى قيمة المعلمة.
 - ج- $E(\bar{X}) = \mu$
 - د- تساوي كلا الوسطين للمجتمع والعينة.
- 3- مجتمع مسوزع توزيعاً طبيعياً انحرافه المعياري $\sigma = 11$ سحبت منه عينة عشوائية حجمها $n = 26$ وكان متوسط العينة $\bar{x} = 40$ احسب تقدير متوسط المجتمع μ بثقة 95%
 - أ- [35.77 , 44.23]
 - ب- [35.77 - 37.23]
 - ج [29 - 51]
 - د- [26 - 37]

4- عينة تتألف من 25 طالب من طلبة الصف السادس وكان متوسط علاماتهم على اختبار في العلوم 80 بالمحرف معياري مقداره 5 ، على اعتبار أن مستوى الثقة هو 95%، جد متوسط أداء طلبة الصف السادس على هذا الاختبار؟

مستوى الثقة	حجم العينة	درجات الحرية	T
95%	30	29	2.462
95%	25	24	2.492
95%	20	19	2.539

ب- [75 - 85]

أ- [74 - 86]

د- [77.508 - 82.492]

ج- [74.508 - 80.492]

5- أخذت عينة تتكون من 100 طالب جامعي من الذكور وسئلوا عن رأيهم في التدخين، فأجاب 60 منهم بالإيجاب، فما هي النسبة العامة لتأييد الذكور للتدخين، إذا كان مستوى الثقة المطلوب هو 99%.

ب- [0.474-0.726]

أ- [0.400-0.700]

د- [0.440 - 0.720]

ج- [0.450-0.702]

6- ما المقصود بعبارة "إن معامل الارتباط بين متغيرين كان ذا دلالة إحصائية"؟

أ- استدلال إحصائي ينتج عن رفض الفرضية الصفرية $H_0: \rho = 0$

ب- تعني أن معامل الارتباط في المجتمع بين المتغيرين لا يختلف عن صفر.

ج- رفض هذه الفرضية الصفرية يعني أن قيمة معامل الارتباط في المجتمع تختلف عن الصفر اختلافاً جوهرياً.

د- جميع ما ذكر صحيح.

7- إذا كان معامل الارتباط بين الطول والوزن عند طلبة المدرسة الثانوية هو 0.62 في حالة عينة أفرادها 67 طالباً وطالبة، فما قيمة معامل الارتباط بين الطول والوزن في هذه الحالة الدراسية بوجه عام إذا كان مستوى الثقة المطلوب هو 0.95

ب- [0.43 - 0.73]

أ- [0.45 - 0.75]

د- [0.54 - 0.57]

ج- [0.47 - 0.77]

س2: أخذت عينة عددها 50 عضو هيئة تدريس في الجامعات الأردنية وأعطوا فحصاً في مهارات الحاسوب فكان متوسطهم 70 بانحراف معياري مقداره 2، أوجد المتوسط والتباين لمستوى المعرفة بمهارات الحاسوب عند أعضاء هيئة التدريس إذا كان مستوى الثقة المطلوب هو 95%.

س3: إذا كان متوسط تحصيل عينة من طلبة الثانوية العامة عددها 100 طالب هو 75 بانحراف معياري مقداره 5، أوجد من ذلك متوسط علامات الثانوية العامة وتباينه في حالة الثانوية العامة إذا كان مستوى الثقة المطلوب هو 95%.

س4: في امتحان كفاءة ضم 500 طالباً، إذا كان عدد الساجحين فيها 300 طالباً، جد من ذلك النسبة العامة للنجاح في مثل هذا النوع من الامتحانات إذا كان مستوى الثقة المطلوب 95%.

س5: إذا كان معامل الارتباط بين المستوى الاقتصادي والذكاء في عينة قوامها 50 فرداً هو 0.8، فجد من ذلك قيمة الارتباط بين هذين المتغيرين بوجه عام إذا كان مستوى الثقة المطلوب 95%.

الفصل الرابع

اختبار الفرضيات

Hypothesis Testing

- 1-4 اختبار الفرضيات Hypothesis Testing
- 2-4 اختبار الفرضيات الإحصائية
- Testing Statistical Hypothesis
- 3-4 مفاهيم أساسية في فحص الفرضيات
- Basic Concepts in Hypothesis Testing
 - * الخطأ من النوع الأول Type One Error
 - * الخطأ من النوع الثاني Type Two Error
 - * مستوى الدلالة Level of Significance
 - * قوة الاختبار Power of the Test
- 4-4 خطوات اختبار الفرضيات Hypothesis Testing Steps
- 5-4 استخدام برمجية SPSS في حل المسائل.
- 6-4 تمارين Exercise

الفصل الرابع

اختبار الفرضيات Hypothesis Testing

1-4 اختبار الفرضيات Hypothesis Testing

أحد أساليب عمل الاستنتاجات الإحصائية والاستدلال الإحصائي والهدف منه الوصول إلى قرار بشأن معلمة المجتمع Parameter من خلال تقديرها من العينة Statistic المأخوذة من ذلك المجتمع.

الفرضية العلمية والفرضية الإحصائية Scientific and Statistical Hypothesis

الفرضية العلمية Scientific Hypothesis

حل مقترح لمشكلة يصاغ بشكل استنتاجي كتخمين ذكي يستند على معلومات علمية سابقة. وتقرر صحة الفرضية العلمية أو خطؤها في ضوء الخبرة والتجربة.

الفرضية الإحصائية Statistical Hypothesis

ادعاء أو تصريح حول معلم غير معلوم تخضع للاختبار الإحصائي الذي يحدد قبولها أو رفضها.

الفرضية الإحصائية التي تخضع للاختبار الإحصائي تسمى بالفرضية الصفرية (H_0).

الفرضية الصفرية (H_0) Null Hypothesis: وتصاغ عادة بالنفي.

H_0 : هي الهدف المطلوب اختباره وقبولها يعني أن العينة متوافقة مع الفرضية.

واختبار الفرضيات هو إجراء إحصائي يستخدمه الباحث لاختبار الفرضية الصفرية ليتبين فيما إذا كانت صائبة أم خاطئة.

أمثلة:

$$H_0: \mu = 70$$

تعني أن متوسط المجتمع لا يختلف عن 70

$$H_0: \sigma^2 = 70$$

تعني أن تباين المجتمع لا يختلف عن 70

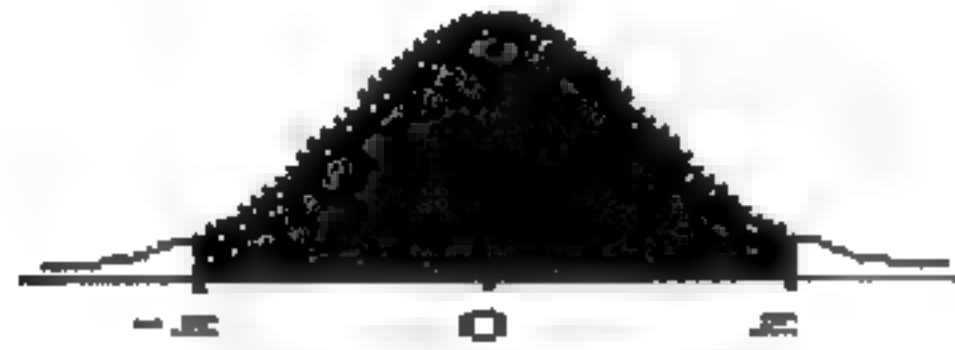
$H_0: P = 0.03$	تعني أن نسبة المجتمع لا تختلف عن 0.03
$H_0: \rho = 0.7$	معامل الارتباط بين المتغيرات في المجتمع لا يختلف عن 0.7
$H_0: \mu_1 - \mu_2 = 0$	إن الفرق بين متوسطي المجتمعين لا يختلف عن صفر
$H_0: \mu_1 = \mu_2$	وسط المجتمع الأول لا يختلف عن وسط المجتمع الثاني
$H_0: \sigma_1^2 - \sigma_2^2 = 0$	إن الفرق بين تباين المجتمعين لا يختلف عن صفر
$H_0: \sigma_1^2 = \sigma_2^2$	تباين المجتمع الأول لا يختلف عن تباين المجتمع الثاني

الفرضية البديلة (H_1) Alternate Hypothesis

تقبل إذا رفضت الفرضية الصفرية وترفض إذا فشلنا في رفض الفرضية الصفرية.

$H_1: \mu \neq 70$	تعني أن وسط المجتمع يختلف عن 70
$H_1: \sigma^2 \neq 70$	تعني أن تباين المجتمع يختلف عن 70
$H_1: P \neq 0.03$	تعني أن نسبة المجتمع تختلف عن 0.03

الفرضية البديلة غير المتجهة Non directional Alternate Hypothesis

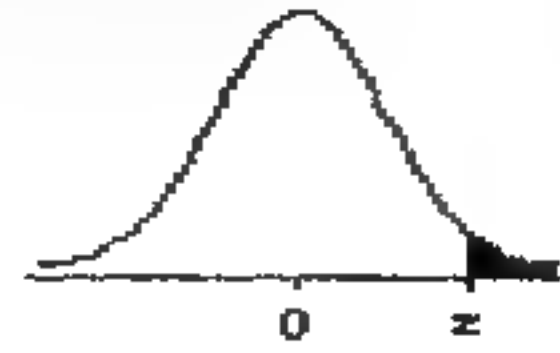
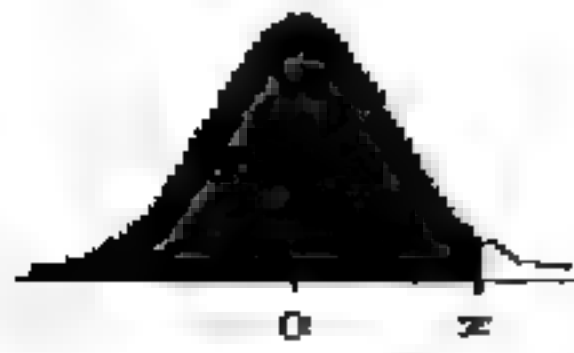


تكون محايدة وتنص على أنه إذا لم تكن للمعلم القيمة المفروضة بالفرضية الصفرية فإن قيمته تختلف عنها بغض النظر عن كون هذه القيمة المقبولة أعلى أو أدنى من القيمة المفروضة بالفرضية الصفرية.

وبعبارة أخرى فإن الفرضية البديلة هي فرضية تقر بوجود الفروق دون أن تحدد اتجاه تلك الفروق.

$H_1: \mu \neq 70$	تعني أن وسط المجتمع يختلف عن 70
$H_1: \mu_1 \neq \mu_2$	وسط المجتمع الأول يختلف عن وسط المجتمع الثاني

الفرضية البديلة المتجهة Directional Alternate Hypothesis



نتم بكون المعلم اكبر أو اصغر من القيمة التي تفترضها الفرضية الصفرية

تعني أن وسط المجتمع اكبر من 70 $H_1: \mu > 70$

تعني أن وسط المجتمع اقل من 70 $H_1: \mu < 70$

وسط المجتمع الأول اكبر من وسط المجتمع الثاني $H_1: \mu_1 > \mu_2$

وعموماً فإن الاختبارات الاحصائية تستخدم البرهان غير المباشر للاستدلال واتخاذ القرار المتعلق بالفرضية البديلة من خلال فحص الفرضية الصفرية التي تسمى ايضاً بالفرضية الاحصائية، أي أن الفرضية البديلة لا تخضع للاختبار الاحصائي والفرضية الصفرية هي التي تخضع للاختبار الاحصائي.

وحيث أن الفرضية تشير دائماً إلى عدم وجود فروق، فإن الفرضية البديلة هي التي تحدد كون الاختبار الاحصائي بذييل واحد أو بذييلين.

فإذا كانت الفرضية البديلة غير متجهة يكون الاختبار الاحصائي بذييلين.

وإذا كانت الفرضية البديلة متجهة يكون الاختبار الاحصائي بذييل واحد.

مثال: يدعي طالب في الثانوية العامة أن معدل عدد ساعات الدراسة اليومية له = 6

ساعات. فإذا اردنا اختبار هذا الإدعاء، فكيف تكتب كل من الفرضية الصفرية والفرضية البديلة.

$$H_0: \mu = 6$$

$$H_1: \mu \neq 6$$

2-4 اختبار الفرضيات الإحصائية

Testing Statistical Hypothesis

إن اتخاذ قرار حول ما إذا كانت الفرضية الصفرية مقبولة أم مرفوضة يتم عن طريق اختبار إحصائي.

الاختبار الإحصائي **Statistical Test**: هو متغير عشوائي ذو توزيع احتمالي يصف العلاقة بين القيم النظرية للمعلم والقيم المحسوبة من العينة. ويتم اتخاذ القرار المتعلق بقبول أو رفض الفرضية الصفرية، وبعد مقارنة قيمة الاختبار الإحصائي المحسوبة من العينة مع قيمته الحرجة المستخرجة من جداول خاصة لتتمكن من اتخاذ القرار.

وعلى أي حال فإن قبول الفرضية الصفرية لا يعني بالضرورة أنها صحيحة إذ أنه يكون ناتجاً عن عدم وجود أدلة كافية من بيانات العينة لرفضها. ولهذا فإن التعبير الأصح في هذه الحالة هو القول بـ "الفشل في رفض الفرضية الصفرية" وليس "قبول الفرضية الصفرية". كما أن رفضها لا يعني أنها خاطئة بل يعني أن الإحصائي كان بعيداً عن المعلم المناظر له في المجتمع.

3-4 مفاهيم أساسية في فحص الفرضيات

Basic Concepts in Hypothesis Testing

لنتذكر أننا في فحص الفرضيات نقوم بتوظيف اختبار إحصائي لفحص فرضية صفرية (هي في حقيقة الأمر إما أن تكون صحيحة وإما أن تكون خاطئة) من أجل أن نصل إلى قرار برفضها أو الفشل في رفضها (قبولها تجاوزاً). وعليه فإننا في هذا السياق نكون في واحدة من الحالات الأربعة التالية:

الفرضية الصفرية			
خاطئة	صحيحة		
قرار صائب	خطأ من النوع الأول	رفض	القرار
خطأ من النوع الثاني	قرار صائب	قبول	

الخطأ من النوع الأول والخطأ من النوع الثاني Type One and Type Two Errors

يظهر نوعان من الخطأ عند اتخاذ القرار حول الفرضية الصفرية:

* **الخطأ من النوع الأول Type One Error**: إتخاذ قرار برفض الفرضية الصفرية

وهي في حقيقة الأمر صحيحة. وهذا قرار خاطيء نكون قد وقعنا فيه.

* **الخطأ من النوع الثاني Type Two Error**: إتخاذ قرار بقبول الفرضية الصفرية وهي

في حقيقة الأمر خاطئة. وهذا قرار خاطيء نكون قد وقعنا فيه.

* إتخاذ قرار برفض الفرضية الصفرية وهي في حقيقة الأمر خاطئة. وهذا قرار صائب لا

غبار عليه.

* إتخاذ قرار بقبول الفرضية الصفرية وهي في حقيقة الأمر صحيحة. وهذا قرار صائب لا

غبار عليه.

مستوى الدلالة وقوة الاختبار Significant Level and Power of the Test

مستوى الدلالة (α) **Significant Level**: الحد الأعلى لاحتمال الوقوع في الخطأ من

النوع الأول، وهي تمثل مساحة الرفض تحت منحنى توزيع اختبار الاحصاءة، وهي احتمال

رفض الفرضية الصفرية عندما تكون صحيحة، وتستخدم القيم 0.01 , 0.05 , 0.10 .

تكون القيمة القصوى لاحتمال ارتكاب الخطأ من النوع الأول (α) وهي مستوى

الدلالة الإحصائية وقيمة مستوى الدلالة α يحددها الباحث لنفسه قبل جمع بياناته من عينة

الدراسة.

فمثلاً $\alpha = 0.05$ تعني إذا تكررت التجربة لعدد كبير جداً فمن المحتمل أن نرفض فرضية

صفرية وهي في الواقع صحيحة 5 مرات في كل 100 مرة، وان الاستنتاج يكون سليماً

وصائباً بثقة 95%، تكون القيمة القصوى لاحتمال ارتكاب الخطأ من النوع الثاني (β).

هناك علاقة بين α ، β فزيادة احدهما يرافقه نقصان الآخر ولكن ليس بنفس المقدار.

* مستوى الثقة $1 - \alpha$

* قوة الاختبار (1 - β) Power of the Test

قدرة الاختبار على رفض الفرضية الصفرية عندما تكون في حقيقة الأمر خاطئة.

$$\text{Power of the test} = 1 - \beta$$

يمكن أن نعين حداً حرجاً للرفض أو القبول للفرضية الصفرية مع الأخذ بالحسبان الفرضية البديلة.

كلما كبرت β يضعف الاختبار، وإن تصغير β يحتاج لتكبير α .

اتخاذ قرارات لاحقة على أساس الدلالة الإحصائية قضية جدلية عند الإحصائيين، فالبعض يرى أن اختبارات الدلالة الإحصائية أفسدت الأبحاث العلمية، لذلك لا بد من التوجه إلى الدلالة العلمية.

$$\beta = \sqrt{n} (\mu - \mu_0) / \sigma$$

* العوامل المؤثرة في قوة الاختبار

1. حجم العينة Sample Size
2. مستوى الدلالة Level of Significance
3. علاقة القيمة الحقيقية للمعلم بقيمته في الفرضية الصفرية.
4. كون الاختبار بذييل أو ذييلين One Tail or Two Tail Test

* اختبار بذييل واختبار بذييلين One and Two Tails of Test

الانحراف عن الفرضية الصفرية باتجاه واحد أو باتجاهين.

$$H_0: \mu \geq \mu_0$$

$$H_0: \mu \leq \mu_0$$

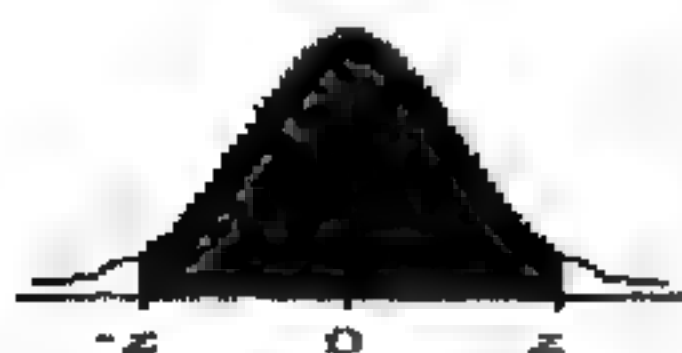
$$H_0: \mu = \mu_0$$

$$H_1: \mu < \mu_0$$

$$H_1: \mu > \mu_0$$

$$H_1: \mu > \mu_0$$

$$H_1: \mu < \mu_0$$



القرار بشأن الاختبار Decision Making

إذا كانت القيمة المحسوبة تقع في منطقة الرفض أي أنها أكبر من القيمة الجدولية نرفض الفرضية الصفرية H_0

إذا كانت القيمة المحسوبة تقع في منطقة القبول أي أنها أقل من القيمة الجدولية نقبل الفرضية الصفرية H_0

4-4 خطوات اختبار الفرضيات Hypothesis Testing Steps

1. تحديد نوع توزيع المجتمع.
 2. صياغة الفرضيتين الصفرية والبديلة.
 3. تحديد مستوى الدلالة (α) المناسب.
 4. تحديد الاختبار الإحصائي المناسب لاختبار الفرضية الصفرية.
 5. إذا كانت P أقل من α نرفض الفرضية الصفرية وبالعكس ذلك نقبل هذه الفرضية.
- * إذا كانت الفرضية البديلة غير متجهة $H_1 : \mu = C$
- مستوى الدلالة يقسم إلى نصفين بالتساوي على كل من ذيلي توزيع المعاينة وتكون منطقة قبول الفرضية في الوسط ومنطقة رفضها على الدليلين.

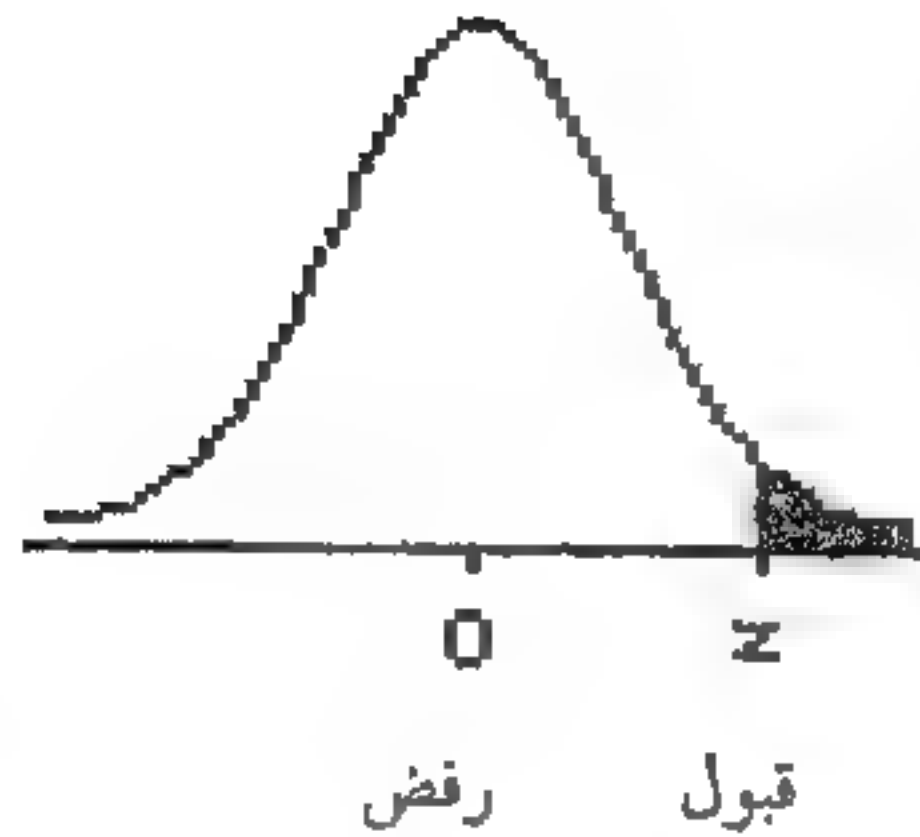
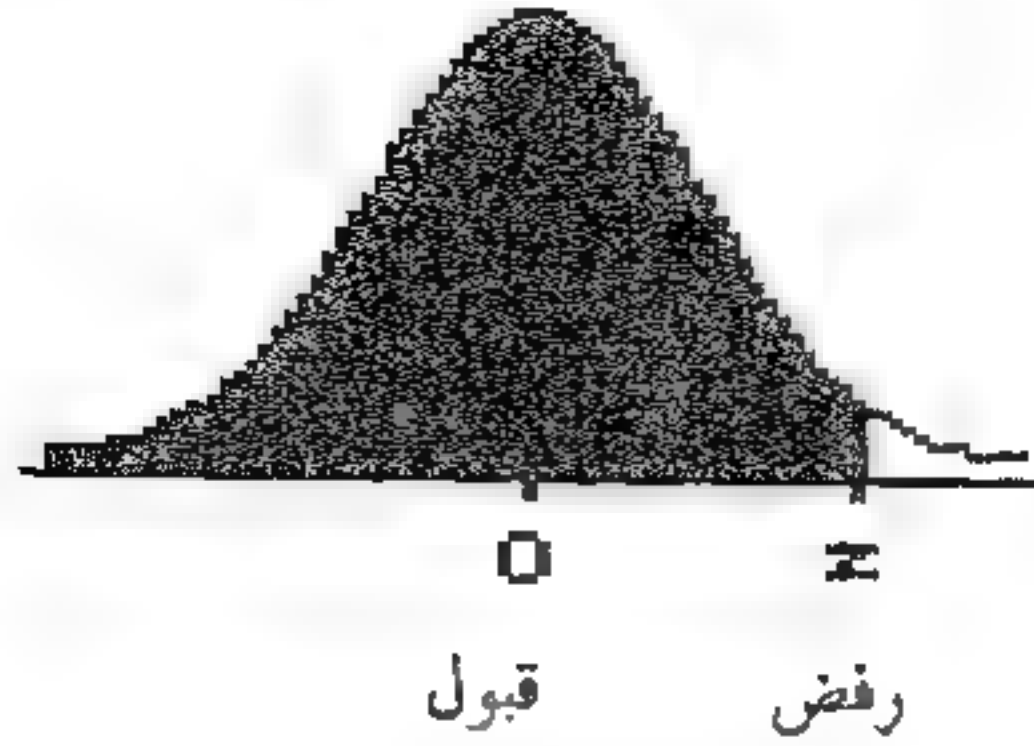


* إذا كانت الفرضية البديلة متجهة

إن منطقة الرفض ستكون إما في الذيل الأيمن أو الأيسر.

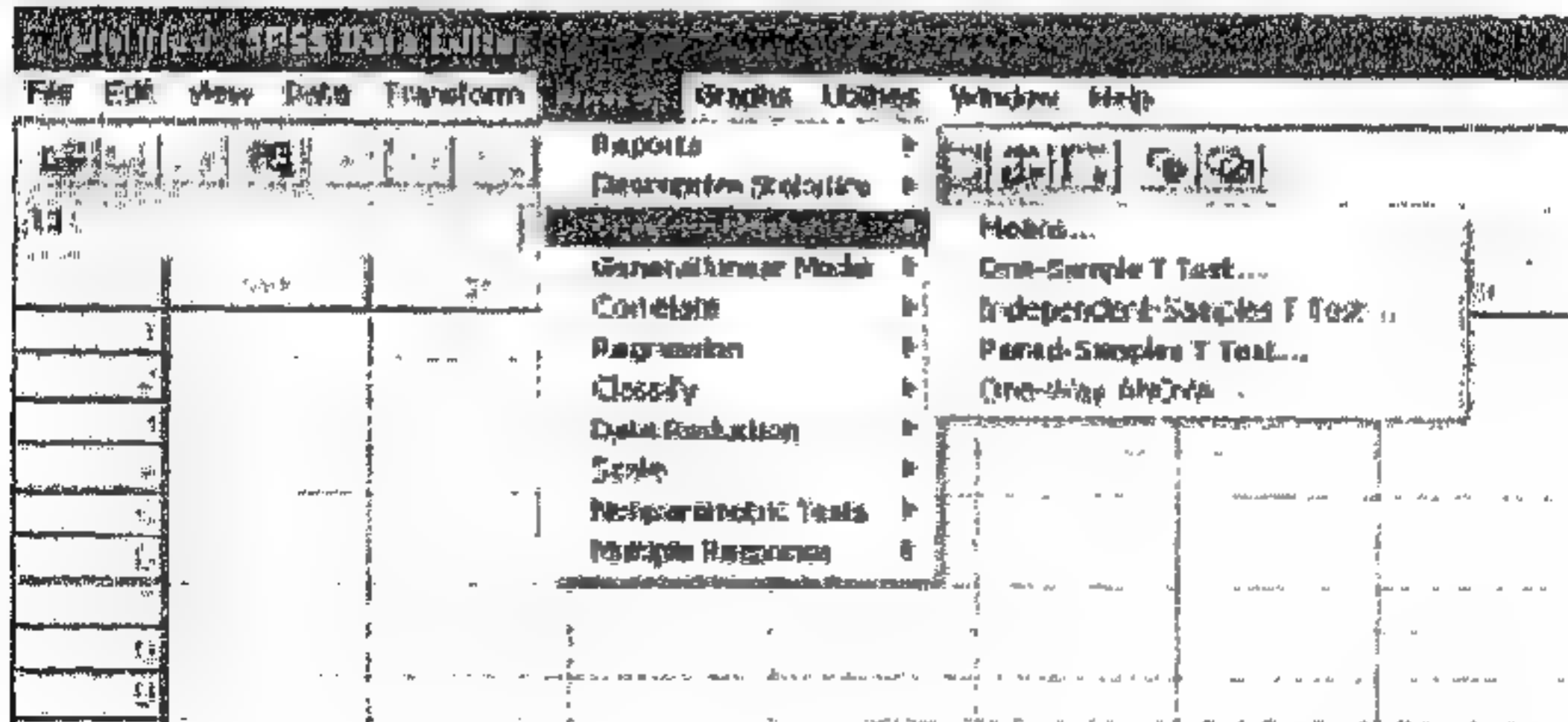
إن منطقة الرفض ستكون في الذيل الأيمن. $H_1 : \mu > C$

إن منطقة الرفض ستكون في الذيل الأيسر. $H_1 : \mu < C$

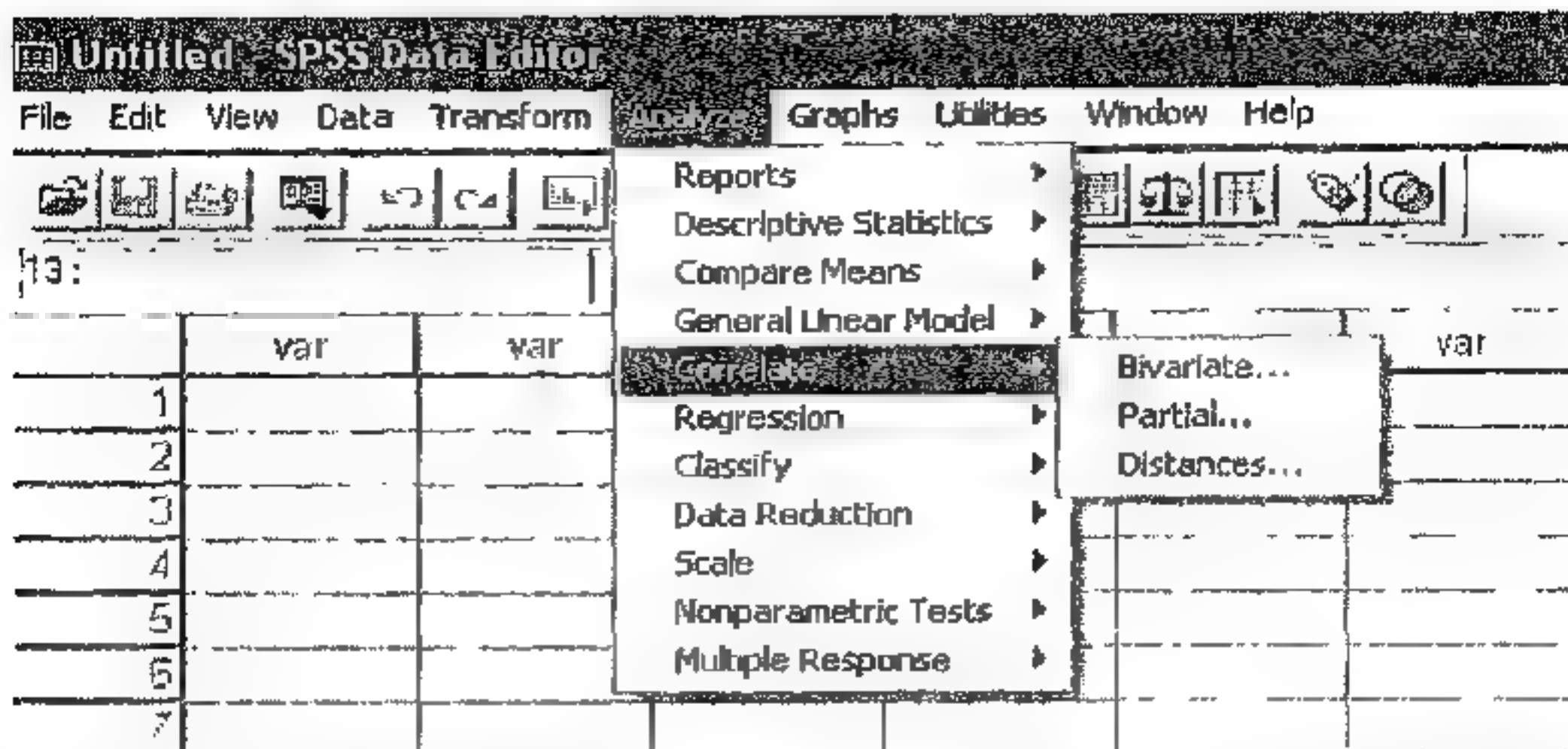


5-4 استخدام برمجية SPSS في حل المسائل.

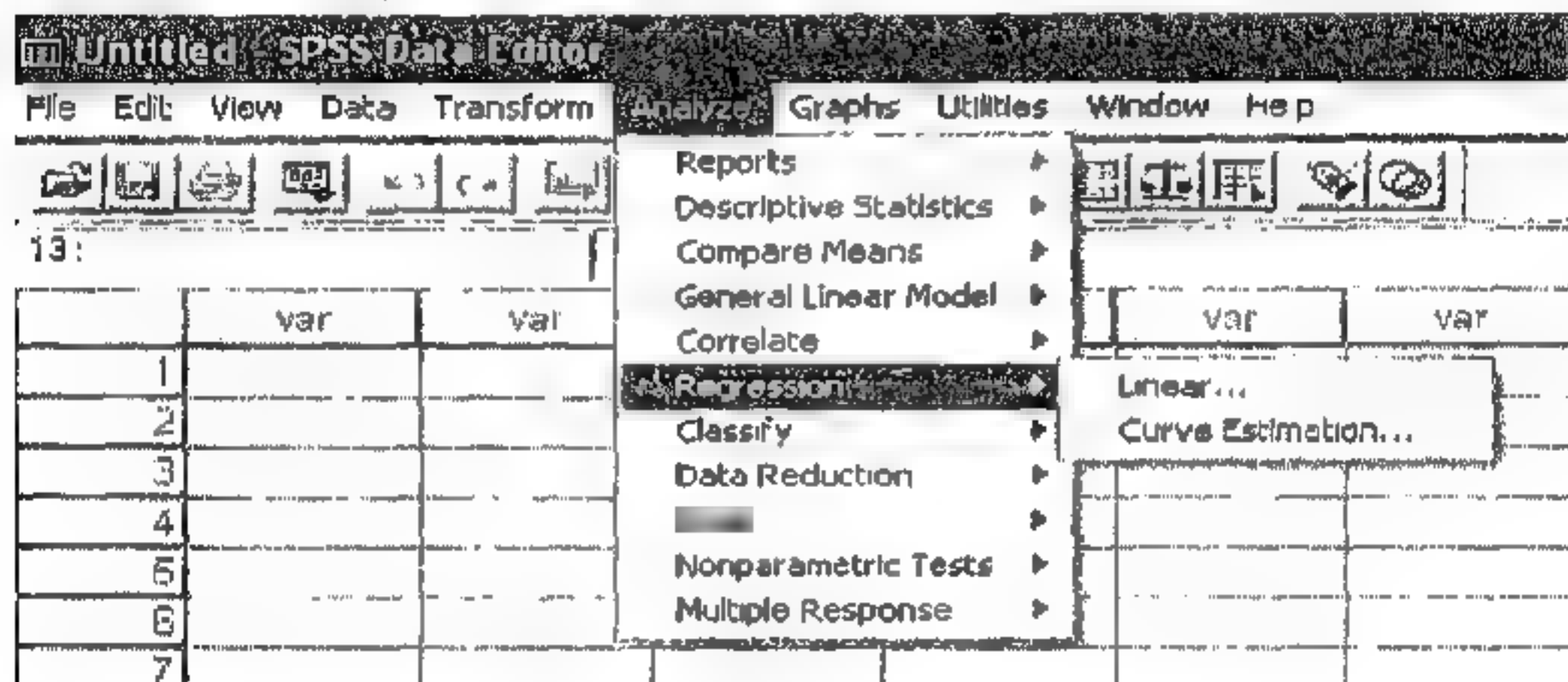
تستخدم برمجية SPSS لاختبار الفرضيات حول الاوساط كالتالي:



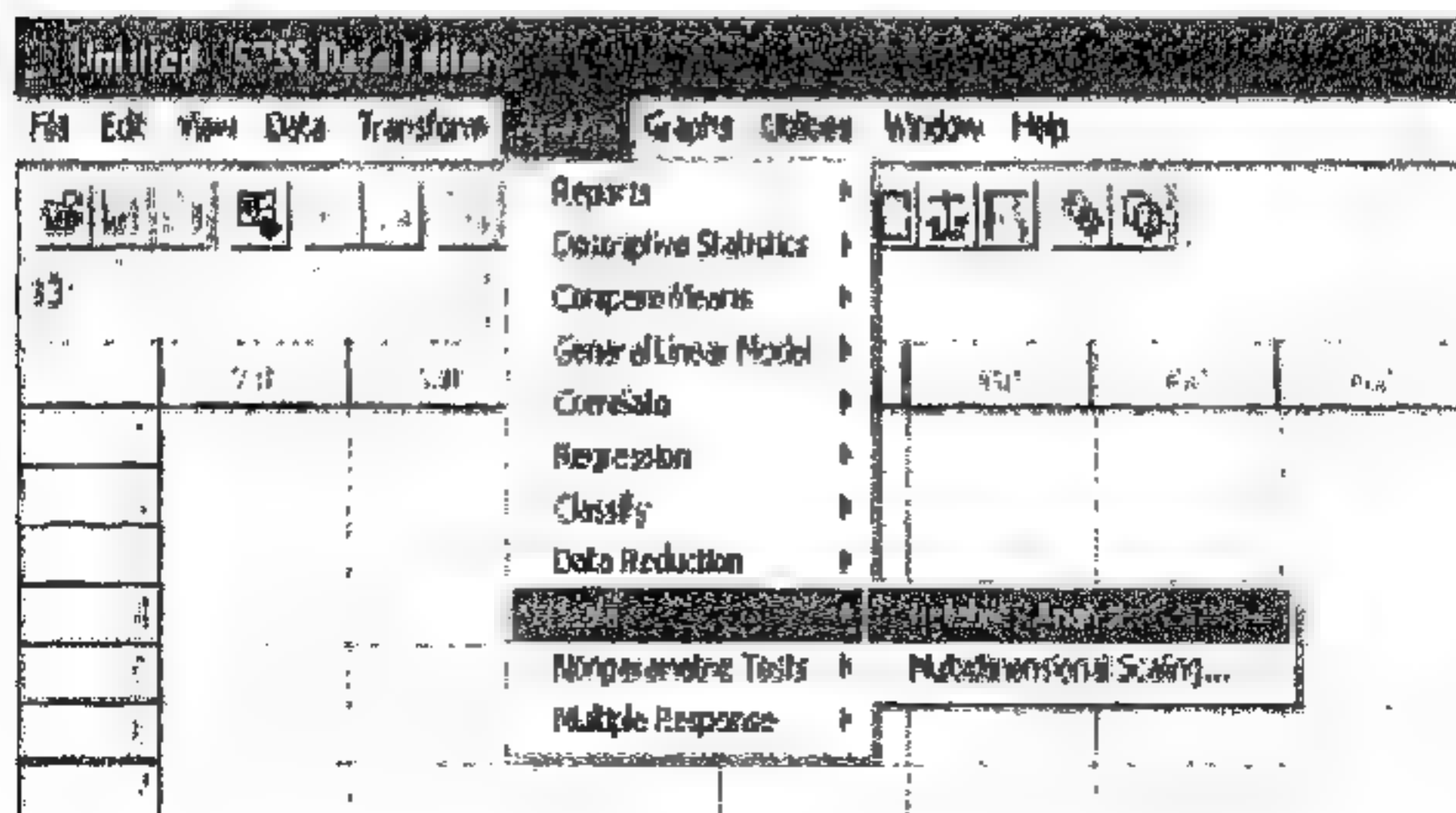
تستخدم برمجية SPSS لاختبار الفرضيات حول معاملات الارتباط كالتالي:



تستخدم برمجية SPSS لاختبار الفرضيات حول الانحدار كالتالي:



تستخدم برمجية SPSS لإيجاد معامل الثبات كالتالي:



كما يمكن اختبار الفرضيات لأشياء كثيرة ستعرف عليها في الفصول القادمة بشيء من التفصيل.

6-4 تمارين Exercise

تمرين (1) : ضع رمز الإجابة (نعم أو لا) في المربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										

- 1- يحقق اختبار الفرضيات والتقدير الإحصائي بفترة الغرض نفسه.
- 2- يستند اختبار الفرضيات على توزيعات المعاينة دائما.
- 3- تتمتع النتائج المنبثقة من اختبار الفرضيات الإحصائية بإمكانية تعميمها وأهميتها حيثما تم الوصول إليها.
- 4- تتمتع نتائج اختبار الفرضية العلمية بالعمومية والأهمية العلمية.
- 5- يتقرر كون الاختبار متجه أو غير متجه من الفرضية الصفرية.
- 6- تقل قوة الاختبار بازدياد حجم العينة.
- 7- الاختبار بذيلين أقل قوة من الاختبار بذيل واحد.
- 8- تصاغ الفرضيات الإحصائية حول المعالم فقط.
- 9- مستوى الدلالة هو الحد الأعلى لاحتمال الوقوع في الخطأ من النوع الأول.
- 10- إذا كانت القيمة المحسوبة تقع في منطقة الرفض أي أنها أقل من القيمة الجدولية نقبل الفرضية الصفرية H_0

تمرين (2) : ضع رمز الإجابة الصحيحة في المربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										

- 1- يتقرر كون الاختبار متجه أو غير متجه من:
 - أ- الفرضية الصفرية.
 - ب- الفرضية البديلة.
 - ج- مستوى الدلالة.
 - د- قوة الاختبار.

2- يتم ارتكاب الخطأ من النوع الأول عند:

- أ- قبول فرضية صفرية خاطئة.
- ب- رفض فرضية صفرية صحيحة.
- ج- رفض فرضية صفرية خاطئة.
- د- قبول فرضية صفرية صحيحة.

3- يتم ارتكاب الخطأ من النوع الثاني عند:

- أ- قبول فرضية صفرية خاطئة.
- ب- رفض فرضية صفرية صحيحة.
- ج- رفض فرضية صفرية خاطئة.
- د- قبول فرضية صفرية صحيحة.

4- قوة الاختبار هي قدرة الاختبار على:

- أ- قبول فرضية صفرية خاطئة.
- ب- رفض فرضية صفرية صحيحة.
- ج- رفض فرضية صفرية خاطئة.
- د- قبول فرضية صفرية صحيحة.

5- تسمى القيمة العظمى لاحتمال ارتكاب الخطأ من النوع الأول ب:

- أ- مستوى القبول.
- ب- مستوى الرفض.
- ج- مستوى الفصل.
- د- مستوى الدلالة.

6- أي من التالية ليست من العوامل المؤثرة في قوة الاختبار:

- أ- حجم العينة.
- ب- كون الاختبار بذيلى أو بذيلىين.
- ج- الإحصائي المستخدم.
- د- مستوى الدلالة.

7- عندما يكون الاختبار بذيلى فإن قوة الاختبار تزداد:

- أ- كلما اقتربت القيمة الحقيقية للمعلم مع القيمة المفروضة.
- ب- كلما ابتعدت القيمة الحقيقية للمعلم عن القيمة المفروضة.
- ب- كلما اقتربت القيمة الحقيقية للمعلم مع القيمة المفروضة.
- ج- كلما تساوت القيمة الحقيقية للمعلم مع القيمة المفروضة.
- د- لا شيء مما ذكر.

8- الذي يحدد كون الاختبار الإحصائي بذيلى أم بذيلىين هو:

- أ- مستوى الثقة.
- ب- مستوى الدلالة.
- ج- الفرضية الصفرية.
- د- الفرضية البديلة.

9- مستوى الدلالة هو:

- أ- الحد الأعلى لاحتمال الوقوع في الخطأ من النوع الأول.
- ب- الحد الأعلى لاحتمال الوقوع في الخطأ من النوع الثاني.
- ج- الحد الأدنى لاحتمال الوقوع في الخطأ من النوع الأول.
- د- الحد الأدنى لاحتمال الوقوع في الخطأ من النوع الأول.

10- إذا كانت القيمة المحسوبة تقع في منطقة الرفض أي أنها أقل من القيمة الجدولية:

- أ- نقبل الفرضية الصفرية H_0
- ب- نقبل الفرضية البديلة H_1
- ج- نرفض الفرضية الصفرية H_0
- د- نرفض الفرضية البديلة H_1

الفصل الخامس

اختبار الفرضيات التي تتعلق

بالمتوسطات الحسابية

Testing Hypothesis Regarding Mean

1-5 اختبار الفرضية المتعلقة بوسط حسابي واحد (مجتمع واحد)

* اختبار الفرضية المتعلقة بوسط حسابي واحد (حجم العينة كبير، تباين المجتمع معلوم)

Testing Hypothesis Regarding Mean (Big sample & σ Known)

* اختبار الفرضية المتعلقة بوسط حسابي واحد (حجم العينة كبير، تباين المجتمع غير معلوم)

Testing Hypothesis Regarding Mean (σ unknown, Big Sample)

* اختبار الفرضية المتعلقة بمتوسط حسابي واحد (العينة صغيرة، تباين المجتمع غير معلوم)

Testing Hypothesis Regarding Mean (σ unknown, Small Sample)

2-5 اختبار الفرضيات حول الفرق بين وسطين حسابيين

* اختبار الفرضيات حول الفرق بين وسطين للبيانات المستقلة

* اختبار الفرضيات حول الفرق بين وسطين للبيانات غير المستقلة

3-5 استخدام برمجية SPSS في حل المسائل.

4-5 تمارين Exercise

الفصل الخامس

اختبار الفرضيات التي تتعلق بالمتوسطات الحسابية

Testing Hypothesis Regarding Mean

1-5 اختبار الفرضيات المتعلقة بوسط حسابي واحد (مجتمع واحد)

اختبار متوسط العينة \bar{X} لمعرفة فيما إذا كان هناك فرق جوهري بين متوسط العينة ومتوسط المجتمع μ المأخوذة منه، أي هل العينة مسحوبة من المجتمع ومثله له، ويمكن أن يستخدم الاختبار الإحصائي Z أو T.

ويستخدم الاختبار الإحصائي T لفحص فرضية تتعلق بالوسط الحسابي وهو اختبار One Sample T-Test ، ويجب أن يتحقق شرطان قبل إجراء الاختبار وهما:

1. يجب أن يكون المتغير موزعاً توزيعاً طبيعياً، ويستعاض عن هذا الشرط بزيادة حجم العينة، وتعتبر العينة كبيرة إذا كان حجمها 30 فأكثر.
2. أن تكون العينة عشوائية وقيم أفرادها لا تعتمد على بعضها البعض.

يستخدم هذا الاختبار لفحص ما إذا كان متوسط متغير ما لعينة واحدة يساوي قيمة ثابتة، وتكتب الفرضية المتعلقة بهذا الاختبار كالتالي:

$$H_0: \mu = a \quad \text{حيث } a \text{ قيمة ثابتة.}$$

ويمكن تحديد قيمة الثابت a كما يلي:

1- العلامة الوسطى على تدرج ما.

- مقياس يتكون من 10 9 8 7 6 5 4 3 2 1 0

المتوسط = 5، لأنها تتوسط مدى الإجابة، يعني $11/55 = 5$ ، وهي قيمة الثابت a

- مقياس ليكرت الخماسي يتكون من 5 4 3 2 1

المتوسط = 3، لأنها تتوسط مدى الإجابة، يعني $5/15 = 3$ ، وهي قيمة الثابت a

الإجابات التي تقل عن قيمة الثابت a تعني فعالية متدنية (سلبية)، الإجابات التي تزيد عن

قيمة الثابت a تعني فعالية عالية (موجبة).

2- من خلال المعلومات السابقة.

مقياس القلق المقنن وله متوسط = 50 درجة.

مقياس الذكاء وله وسط حسابي = 100

الاختبارات المقننة.

3- عدد الإجابات الصحيحة بطريقة الصدفة في امتحان ما.

اختبار مكون من 40 فقرة لكل منها 4 بدائل

احتمال إجابة أي سؤال بطريقة عشوائية = 0.25

عدد الإجابات المتوقع أن تكون صحيحة على جميع الاختبار = $0.25 \times 40 = 10$ درجة

إذا قل متوسط الإجابات عن 10 يكون الاختبار صعب، وإذا زاد عنها فإننا نرفض

فرضية الباحث وهي أن الاختبار ملائم.

1- اختبار الفرضية المتعلقة بوسط حسابي واحد (حجم العينة كبير، تباين المجتمع معلوم)

Testing Hypothesis Regarding Mean (Big sample & σ Known)

بما أن توزيع المعاينة يقترب من التوزيع الطبيعي عندما يصبح حجم العينة أكبر من 30 ،

وتباين المجتمع معلوم، لذلك نستعمل اختبار (Z).

$$Z = \frac{X' - \mu}{\sigma_x / \sqrt{n}}$$

مثال (5-1) يعتقد مدير مدرسة الهمزة الثانوية أن متوسط أداء طلابه في امتحان الثانوية

العامة لا يختلف عن المتوسط العام لطلاب المملكة. فإذا علم أن عدد طلاب مدرسته يساوي

$n = 81$ وأن متوسط معدلاتهم كان $X' = 58$ وأن متوسط المعدلات لطلاب المملكة كان 60

$\mu =$ بانحراف معياري يساوي $\sigma_x = 8$ ، فهل تدعم هذه البيانات ما يعتقد المدير؟ استخدم

$$\alpha = 0.05$$

* خطوات اختبار الفرضيات

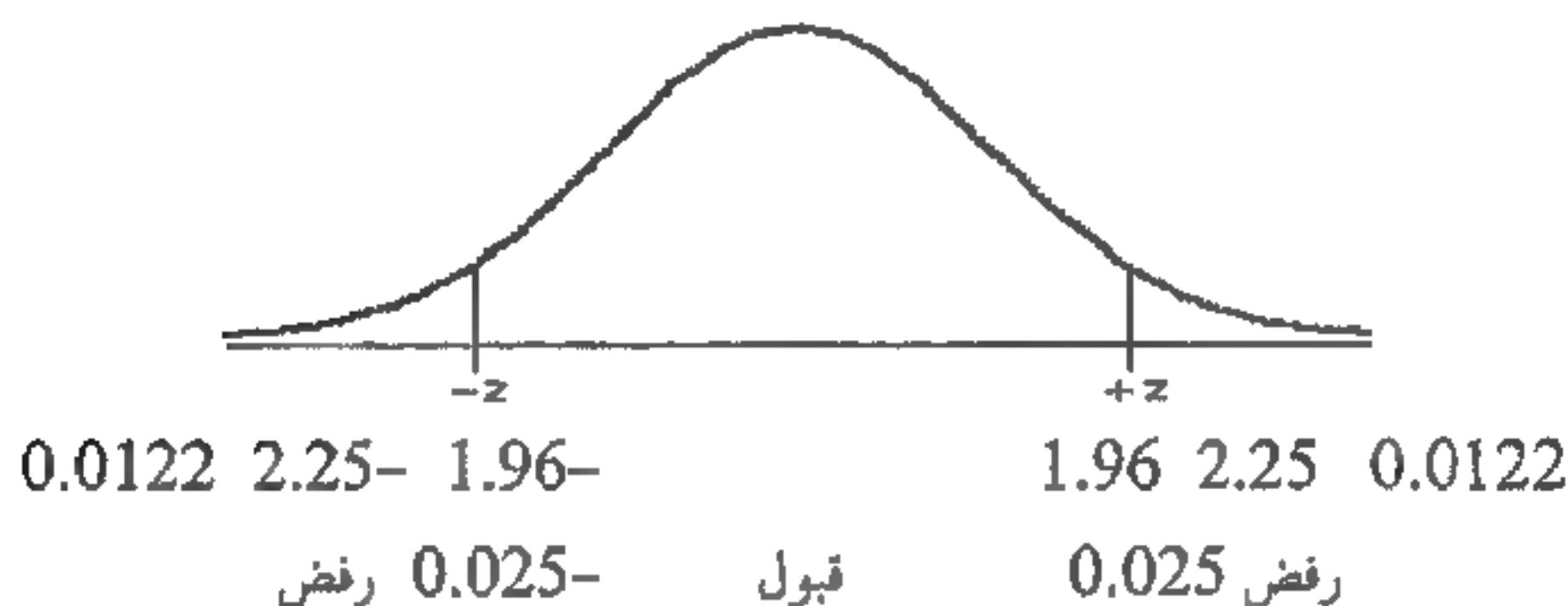
1- تحديد نوع توزيع المجتمع.

توزيع المعاينة يقترب من التوزيع الطبيعي عندما يصبح حجم العينة أكبر من 30 ، لذلك نستعمل اختبار (Z).

2- صياغة الفرضيتين الصفرية والبدلية.

تعني أن وسط المجتمع لا يختلف عن 60
 $H_0: \mu = 60$
 تعني أن وسط المجتمع يختلف عن 60
 $H_1: \mu \neq 60$

3- تحديد مستوى الدلالة (α) المناسب. الفرضية البديلة، استخدم $\alpha = 0.05$



4- تحديد الاختبار الإحصائي المناسب لاختبار الفرضية الصفرية.

بما أن توزيع المعاينة يقترب من التوزيع الطبيعي عندما يصبح حجم العينة أكبر من 30 ، لذلك نستعمل اختبار (Z).

$$Z = \frac{X' - \mu}{\frac{\sigma_x}{\sqrt{n}}} = \frac{58-60}{\frac{8}{\sqrt{81}}} = \frac{-2}{\frac{8}{9}} = -2.25$$

5- إذا كانت (القيمة المحسوبة) أكبر من < القيمة الجدولية α نرفض الفرضية الصفرية وبعكس ذلك نقبل الفرضية الصفرية.

إذا كانت قيمة إحصاء الاختبار ضمن منطقة الرفض يتم رفض الفرضية الصفرية H_0 وغير ذلك لا نستطيع رفضها ونقبل الفرضية البديلة H_1 .

من الجدول: المساحة الواقعة دون -2.25 تساوي 0.0122

بما أن القيمة المحسوبة ($z=2.25$) < 0.0122 القيمة الجدولية

($z=.96$ $\alpha/2=0.05/2= 0.025$) نرفض الفرضية الصفرية ونقبل البديلة، وهذا يعني أن

أداء طلاب مدرسة الهمة في امتحان الثانوية العامة يختلف عن المتوسط العام لطلاب المملكة.

6- عن طريق الحدود الحرجة

الحدود الحرجة عند استخدام اختبار Z ومستوى الدلالة $\alpha = 0.05$ و $\alpha = 0.01$

نوع الاختبار/مستوى الدلالة	$\alpha = 0.05$	$\alpha = 0.01$
بذيلين	$1.96 \pm$	$2.58 \pm$
بذيل	$1.64 \pm$	$2.34 \pm$

بما أن $Z = -2.25$ وهي تقع في منطقة الرفض إذا نرفض الفرضية الصفرية ونقبل

الفرضية البديلة.



منطقة رفض 1.96 منطقة قبول -1.96 -2.25 منطقة رفض

مثال (5-2) إذا كان متوسط علامات الطلاب في مادة الإحصاء هو 70 بالانحراف

معياري مقداره 10 درجات وسحبت عينة عشوائية حجمها 50 طالباً وكان معدل علاماتهم

75، هل هناك فروق جوهرية بين العلامات بمستوى دلالة $\alpha = 0.05$ ؟

الحل: حجم العينة كبير وتباين المجتمع معلوم لذلك نستخدم احصاءة Z وهي:

$$Z = \frac{\bar{X} - \mu}{\sigma_x / \sqrt{n}} = \frac{75 - 70}{10 / \sqrt{50}} = \frac{5}{10 / 7.07} = 3.54$$



3.54 منطقة رفض 1.96 منطقة قبول 1.96- منطقة رفض

القرار: بما أن قيمة Z المحسوبة = 3.54 > قيمة $Z_{\alpha/2}$ الجدولية = 1.96

بما أن قيمة Z المحسوبة (3.54) > أكبر من قيمة $Z_{\alpha/2}$ الجدولية (1.96) وهي تقع في منطقة الرفض للفرضية الصفرية، لذلك نرفض H_0 وهذا يعني أن مستوى طلبة العينة يختلف عن المستوى العام بدرجة ثقة 95%.

2- اختبار الفرضية المتعلقة بوسط حسابي واحد (حجم العينة كبير، تباين المجتمع غير معلوم)

Testing Hypothesis Regarding Mean (σ unknown, Big Sample)

بما أن توزيع المعاينة يقترب من التوزيع الطبيعي عندما يصبح حجم العينة أكبر من 30 ، وتباين المجتمع غير معلوم، لذلك نستعمل اختبار (Z).

$$Z = \frac{X' - \mu}{S_x / \sqrt{n}}$$

مثال 5-3 إذا كانت أوزان الطلاب في المدرسة تخضع لتوزيع طبيعي وسطه 60 كغم، وسحبت منه عينة عشوائية حجمها 50 طالباً وسجلت متوسطاً قدره 65 كغم بانحراف معياري قدره 10 كغم، اختبر الفرضية أن المتوسط أكبر من 60 باستخدام $\alpha = 0.05$ ؟

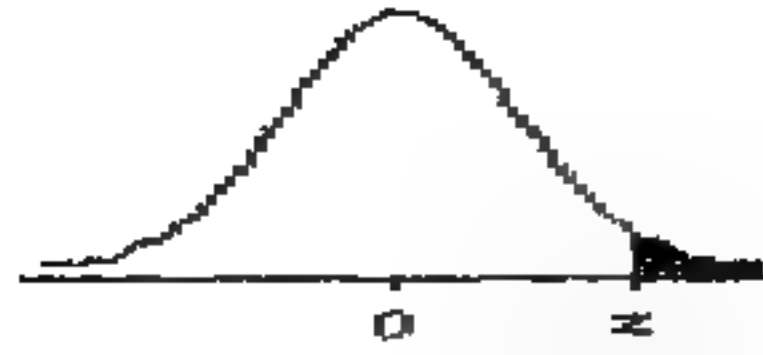
$H_0: \mu = 60$ تعني أن وسط المجتمع لا يختلف عن 60

$H_1: \mu > 60$ تعني أن وسط المجتمع أكبر من 60

$$Z = \frac{X' - \mu}{S_x / \sqrt{n}} = \frac{65 - 60}{10 / \sqrt{50}} = \frac{5}{10 / 7.07} = 3.54$$

بما أن الاختبار متجه أي بذييل واحد فإن مستوى الدلالة هو $\alpha = 0.05$ والقيمة

الجدولية $Z_{\alpha} = 1.645$



منطقة الرفض 1.64 منطقة عدم الرفض

القرار:

بما أن قيمة Z المحسوبة (3.54) < أكبر من قيمة Z_{α} الجدولية (1.64) وهي تقع في منطقة الرفض لذلك نرفض H_0 ونقبل الفرضية البديلة وهذا يعني أن معدل وزن العينة أكبر من 60 وهو المستوى العام بدرجة ثقة 95%.

* إذا كانت الفرضية غير متجهة

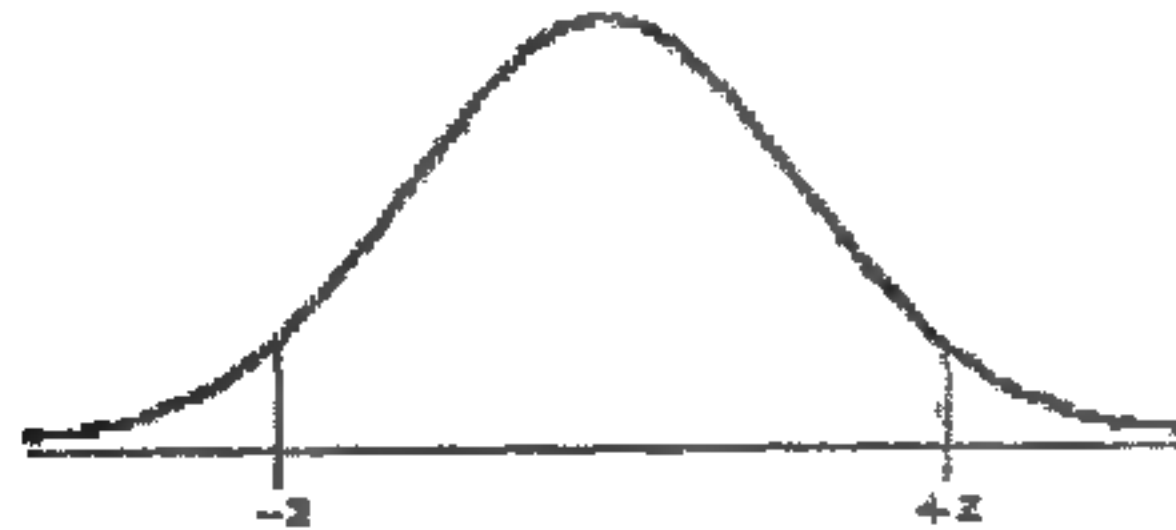
$H_0: \mu = 60$ تعني أن وسط المجتمع لا يختلف عن 60

$H_1: \mu \neq 60$ تعني أن وسط المجتمع يختلف عن 60

$$Z = \frac{X' - \mu}{S_x / \sqrt{n}} = \frac{65 - 60}{10 / \sqrt{50}} = \frac{5}{10 / 7.07} = 3.54$$

بما أن الاختبار غير متجه أي بذيلين فإن مستوى الدلالة هو $\alpha = 0.05 / 2$

0.025 والقيمة الجدولية $Z_{\alpha/2} = 1.96$



منطقة رفض 1.96 منطقة قبول -1.96 منطقة رفض

القرار:

بما أن قيمة Z المحسوبة (3.54) < أكبر من قيمة Z_{α} الجدولية (1.96) وهي تقع في منطقة الرفض لذلك نرفض H_0 ونقبل الفرضية البديلة وهذا يعني أن معدل وزن العينة يختلف عن 60.

3- اختبار الفرضية المتعلقة بمتوسط حسابي واحد (σ غير معلومة والعينة صغيرة الحجم)

Testing Hypothesis Regarding Mean (σ unknown, Small Sample)

عندما يكون الانحراف المعياري لأداء المجتمع σ غير معلوم، والعينة صغيرة الحجم، فإن الاختبار المناسب هو اختبار T ، وفي هذه الحالة نقدر الخطأ المعياري لتوزيع المعاينة من خلال العلاقة

S_x انحراف معياري عينة

$$\sigma_{x'} = S_{x'} = \frac{S_x}{\sqrt{n}}$$

$$T = \frac{X' - \mu}{S_x / \sqrt{n}}$$

تقرر الحدود الحرجة للرفض أو القبول في ضوء درجات الحرية γ ومستوى لدلالة α ، وكون الفرضية البديلة متجهة أو غير متجهة.

مثال 5-4

- 1- استخدم برنامج لتعليم اللغة الإنجليزية مع 16 طالباً من طلاب جامعة الإسراء، وقد اعتقد مصممو البرنامج أنه يرفع متوسط أداء الطلاب إلى أكثر من 0.70 وعند إنهاء الطلاب لبرنامج أجري اختبار مقنن فكان متوسط أدائهم = 72.6 بانحراف معياري غير متحيز = 5 ، هل تدعم هذه النتائج صحة ادعاء مصممي البرنامج؟ استخدم $\alpha = 0.05$
- 2- وإذا كان ادعاء مصممي البرنامج أن برنامجهم يجعل علامات الطلاب الضعاف لا تختلف عن المتوسط العام وهو 70 ، فهل تدعم البيانات صحة ادعاء مصممي البرنامج؟ استخدم $\alpha = 0.05$

1- حل الحالة الأولى:

$$H_0: \mu > 70$$

$$H_1: \mu \leq 70$$

عندما يكون الانحراف المعياري لأداء المجتمع σ غير معلوم، فإن الاختبار المناسب T

نحدد القيم الحرجة للرفض أو القبول:

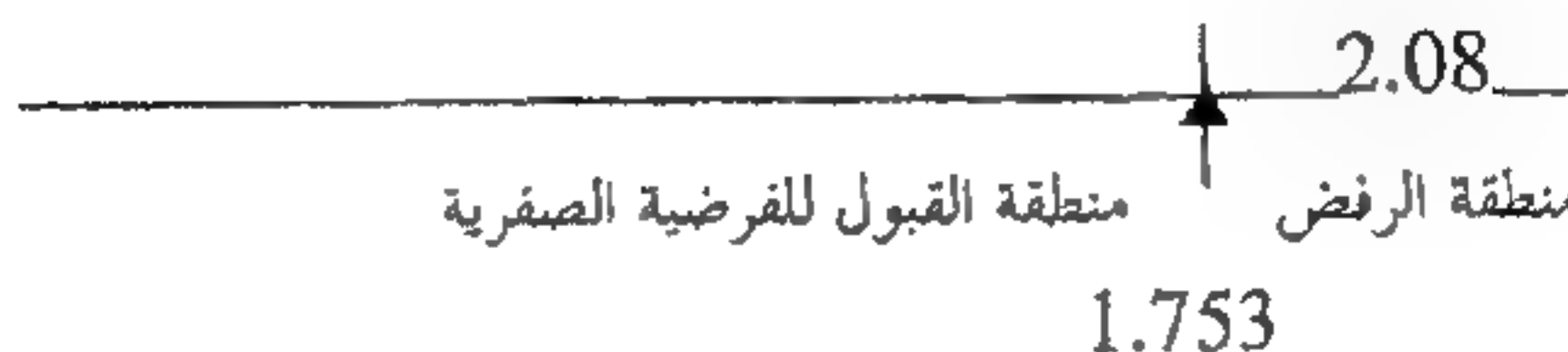
بما أن مستوى الدلالة $\alpha = 0.05$

درجة الحرية $\gamma = n-1 = 16 - 1 = 15$

الاختبار بذييل واحد

بالاستعانة بالجدول رقم 3 تكون القيمة الحرجة $T_{15,0.95} = 1.753$

تكون منطقتا القبول والرفض كما هو مبين أدناه



الآن نحسب قيمة T كما يلي

$$T = \frac{X' - \mu}{\frac{Sx}{\sqrt{n}}} = \frac{72.6 - 70}{\frac{5}{\sqrt{16}}} = 2.08$$

بما أن قيمة T المحسوبة (2.08) تقع في منطقة الرفض للفرضية الصفرية، لذلك نرفض الفرضية الصفرية ونقبل البديله ونقول أن متوسط أداء الطلاب الذين خضعوا للبرنامج لم يكن أعلى من 70 بمستوى الدلالة $\alpha = 0.05$

2- وإذا كان ادعاء مصممي البرنامج أن برنامجهم يجعل علامات الطلاب الضعاف لا تختلف عن المتوسط العام وهو 70 ، فهل تدعم البيانات صحة ادعاء مصممي البرنامج؟
استخدم $\alpha = 0.05$

$$H_0: \mu = 70$$

$$H_1: \mu \neq 70$$

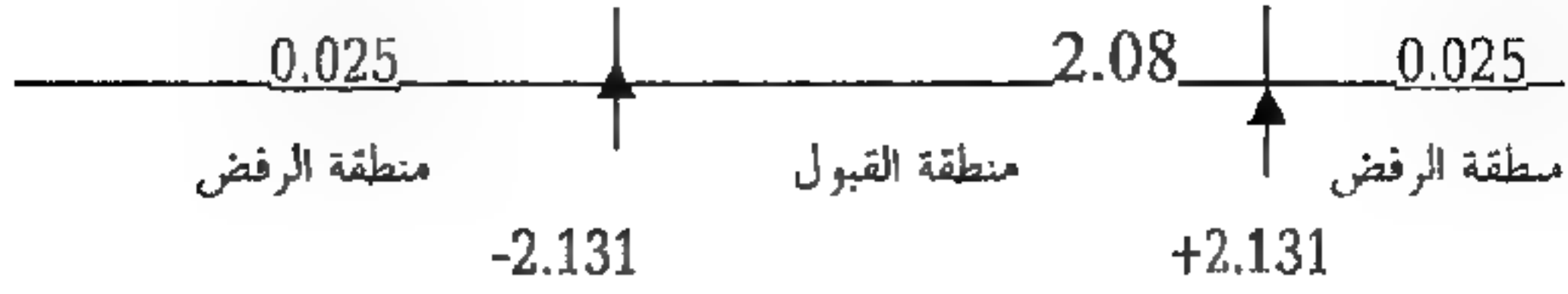
بما أن مستوى الدلالة $\alpha = 0.05$ والاختبار بذييلين

درجة الحرية $\gamma = n-1 = 16 - 1 = 15$

درجة الحرية $\gamma = n-1 = 16 - 1 = 15$

بالاستعانة بجدول توزيع t تكون القيمة الحرجة $T_{15,0.95} = \pm 2.131$

تكون منطقتا القبول والرفض كما هو مبين أدناه



الآن نحسب قيمة T كما يلي

$$T = \frac{X' - \mu}{\frac{Sx}{\sqrt{n}}} = \frac{70 - 72.6}{\frac{5}{\sqrt{16}}} = 2.08$$

بما أن قيمة T المحسوبة (2.08) تقع في منطقة القبول للفرضية الصفرية، لذلك نقبل الفرضية الصفرية ونقول أن متوسط أداء الطلاب الذين خضعوا للبرنامج لم يختلف عن المتوسط العام وهو 70. بمستوى الدلالة $\alpha = 0.05$

2-5 اختبار الفرضيات حول الفرق بين وسطين حسابيين Testing Hypothesis Regarding the Difference between Two Means

تستخدم لإجراء المقارنة بين وسطين لمجتمعين في حالة البيانات المستقلة والبيانات غير المستقلة.

1- اختبار الفرضيات حول الفرق بين وسطين حسابيين (للبيانات المستقلة).

البيانات المستقلة: البيانات التي لا يوجد فيما بينها ارتباط. أداء مجموعة للذكور وأخرى من الإناث.

X'_1 وسط العينة الأولى مستقلاً عن X'_2 وسط العينة الثانية، الفرضية غير متجهة

$$H_0: \mu_1 - \mu_2 = 0$$

$$H_1: \mu_1 - \mu_2 \neq 0$$

الفرضية عندما تكون متجهة

$$H_0: \mu_1 - \mu_2 > 0$$

$$H_1: \mu_1 - \mu_2 < 0$$

يتخذ التوزيع شكل T بدرجات حرية $\gamma = n_1 + n_2 - 2$

n_1 : حجم العينة الأولى.

n_2 : حجم العينة الثانية.

الاختبار الإحصائي المناسب:

$$T = \frac{\bar{X}_1 - \bar{X}_2}{s_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

الافتراضات التي يجب توفرها لتطبيق اختبار T

1. التوزيع الطبيعي Normality

يجب أن يكون توزيع متغير الاختبار طبيعياً في كل فئة من فئات متغير التجميع، وإذا كانت العينة 30 فأكثر يمكن الاستغناء عن هذا الشرط.

2. تجانس التباين في المجتمع Homogeneity

يجب أن يكون تباين متغير الاختبار متساوياً في كلا فئتي متغير التجميع.

3. العشوائية، والاستقلالية Independence

مثال (5-5) أراد باحثاً أن يعرف فعالية أسلوب معين في التدريس لطلبة الجامعة في مستوى السنة الأولى، فأخذ شعبتين كل شعبة تتكون من 25 طالب عشوائياً في مجموعتين ثم عين عشوائياً أحدهما لتكون مجموعة تجريبية والأخرى ضابطة. وفي نهاية التجربة أعطيت المجموعتين اختباراً موحداً فكانت النتائج كما يلي

المجموعة الضابطة	المجموعة التجريبية
$n_2 = 25$	$n_1 = 25$
$\bar{X}_2 = 6$	$\bar{X}_1 = 7.65$
$S_{x_2} = 2.43$	$S_{x_1} = 2.55$

هل تدل هذه البيانات على أن أداء المجموعة التجريبية كان أفضل من أداء المجموعة

الضابطة على مستوى دلالة $\alpha = 0.05$

$$H_0: \mu_1 = \mu_2$$

الفرضية الصفرية

$$H_1: \mu_1 > \mu_2$$

الفرضية عندما تكون متجهة

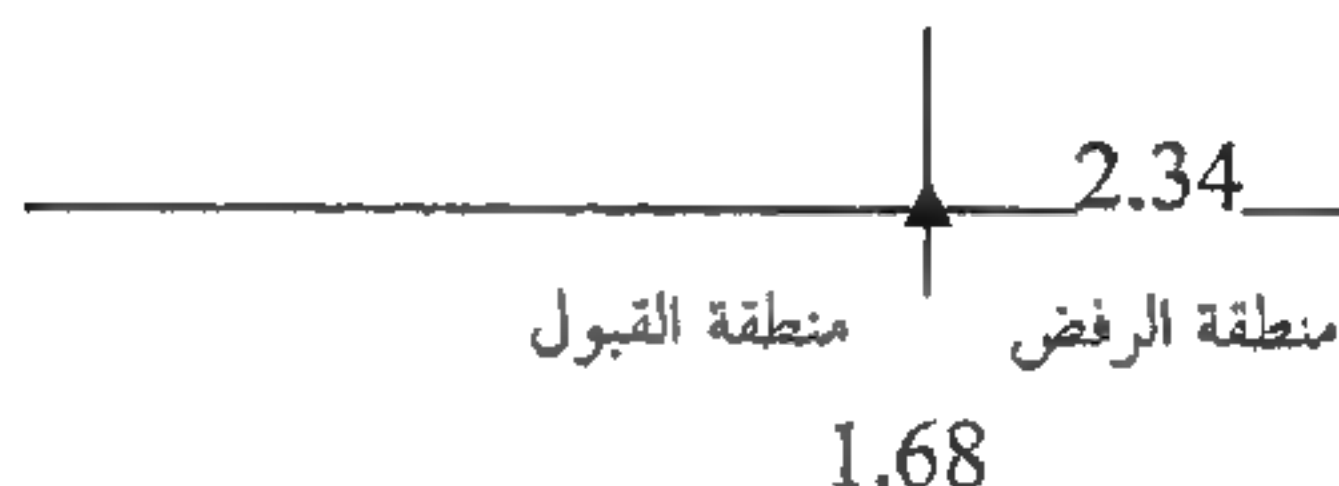
نحدد القيمة الحرجة للرفض أو القبول:

بما أن مستوى الدلالة $\alpha = 0.05$ ، والاختبار بذييل واحد.

$$\gamma = 25 + 25 - 2 = 48 \text{ درجات الحرية}$$

تكون قيمة T الحرجة من الجدول $T_{48,0.05} = 1.68$

$$\alpha = 0.05$$



نحسب الخطأ المعياري لتوزيع المعاينة كما يلي:

$$S_d^2 = \frac{(n_1 - 1) S_1^2 + (n_2 - 1) S_2^2}{n_1 + n_2 - 2} = \frac{24 * 2.55^2 + 24 * 2.43^2}{25 + 25 - 2} = 6.204$$

$$S_d = 2.491$$

نحسب قيمة الإحصائي T كما يلي:

$$T = \frac{\bar{X}_1 - \bar{X}_2}{S_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$T = \frac{X'_1 - X'_2}{S_d \sqrt{1/n_1 + 1/n_2}} = \frac{7.65 - 6}{2.491 \sqrt{1/25 + 1/25}} = 2.34$$

نقارن القيمة المحسوبة (2.34) بالقيمة الحرجة (1.68) فنجد أنها تقع في منطقة الرفض.

إن البيانات تدل على أن الذين يخضعون للبرنامج التدريبي يصبح أداؤهم أفضل من الذين

لا يخضعون على مستوى دلالة $\alpha = 0.05$

تبعات عدم الإيفاء بافتراضات اختبار T

1- التوزيع الطبيعي Normality

2- تجانس التباين في المجتمعين Homogeneity

3- الاستقلالية Independence

1- افتراض التوزيع الطبيعي Normality

إن هذا الافتراض يمكن مخالفته بدون تبعات تذكر.

2- افتراض الاستقلالية Independence

يهدد البحث بأكمله وعلى الباحث أن لا يتهاون في ذلك إطلاقاً.

3- تجانس التباين في المجتمعين Homogeneity

■ يمكن مخالفته إذا تساوت العينتان في أعداد أفرادهما. إذا كانت $n_1 = n_2$

■ ولكن عندما تكون $n_1 \neq n_2$ فهناك وضعين:

1- العينة الكبيرة الحجم منتمية للمجتمع ذي التباين الكبير

والعينة الصغيرة الحجم منتمية للمجتمع ذي التباين الصغير

فإن احتمال ارتكاب الخطأ من النوع الأول تكون أقل من α إذاً يكون الباحث في

الجانب الآمن.

2- العينة الكبيرة الحجم منتمية للمجتمع ذي التباين الصغير

والعينة الصغيرة الحجم منتمية للمجتمع ذي التباين الكبير

فإن احتمال ارتكاب الخطأ من النوع الأول تزداد.

2- اختبار الفرضيات حول الفرق بين وسطين حسابيين (للبينات غير المستقلة).

البيانات غير المستقلة: البيانات التي يوجد فيما بينها ارتباط.

مثال: إجراء اختبار على نفس المجموعة مرتين، $n_1 = n_2 = n$

مثال: عينة من الأزواج وزوجاتهم.

X'_1 وسط العينة الأولى مرتبطاً مع X'_2 وسط العينة الثانية وهي نفس العينة الأولى.

اختبار T للعينات المترابطة Paired Sample T-Test: يستخدم لفحص فرضية متعلقة

بمساواة متوسط متغيرين أو مساواة متوسط متغير لعيتين غير مستقلتين.

لضمان دقة نتائج اختبار T يجب تحقق الشروط التالية:

1. أن يكون توزيع الفرق بين المتغيرين طبيعياً. ويمكن تجاهله إذا كان حجم العينة أكثر

من 30

2. أن تكون العينة عشوائية.

الفرضية عندما تكون غير متجهة

$$H_0: \mu_1 - \mu_2 = 0$$

$$H_1: \mu_1 - \mu_2 \neq 0$$

الفرضية عندما تكون متجهة

$$H_0: \mu_1 - \mu_2 > 0$$

$$H_1: \mu_1 - \mu_2 < 0$$

يتخذ التوزيع شكل T بدرجات حرية $\gamma = n - 1$

n: عدد أزواج المشاهدات. وسط التوزيع = 0

الانحراف المعياري، الخطأ المعياري لتوزيع المعاينة لفروق الأوساط $Sx'_1 - x'_2$

$$Sx'_1 - x'_2 = \sqrt{S^2 x'_1 + S^2 x'_2 - 2d S x'_1 S x'_2}$$

$$S^2 x'_1 = \frac{S^2_1}{n}$$

$$S^2 x'_2 = \frac{S^2_2}{n}$$

$$S = \sqrt{\frac{\sum x^2 - (\sum x)^2}{n - 1}}$$

مثال 5-6) إذا كانت علامات طلاب شعبة في مادة الإحصاء في الاختبار الأول

والاختبار الثاني كما هو مبين أدناه، والمطلوب اختبار الفرق بين متوسطي درجات الاختبارين

في مادة الإحصاء.

الفرضية:

$$H_0: \mu_1 - \mu_2 = 0$$

$$H_1: \mu_1 - \mu_2 \neq 0$$

$$d_i = X_i - Y_i$$

1- لحسب معدل الفروق ما بين العينتين:

$$d' = \frac{\sum d_i}{n}$$

2- نحسب تباين الفروق:

$$S_d^2 = \frac{\sum d_i^2 - (\sum d')^2}{n - 1}$$

3- نستخدم الاختبار t كما يلي:

$$T = \frac{\bar{X} - \mu_a}{\frac{S_a}{\sqrt{n}}}$$

Stud	Stat1_m	Stat2_m	di=Stat1-Stat2	di ²
1	20	15	5	25
2	20	14	6	36
3	18	20	-2	4
4	15	17	-2	4
5	16	17	-1	1
6	12	18	-6	36
7	15	20	-5	25
8	18	18	0	0
9	18	14	4	16
10	18	17	1	1
11	14	8	6	36
12	16	18	-2	4
13	12	14	-2	4
14	17	14	3	9
15	14	12	2	4
16	19	20	-1	1
17	11	10	1	1
18	10	10	0	0
19	19	20	-1	1
20	12	15	-3	9
21	17	20	-3	9
المجموع			0	226
d'			0	

$$S_d^2 = \frac{d_i^2 - (\frac{d'}{n})^2}{n - 1} = 11.3$$

$$S_d = 3.361547$$

$$T = \frac{\bar{X} - \mu_a}{\frac{S_d}{\sqrt{n}}}$$

المحسوبة $T = (0-0) / (3.36/4.58) = 0$

$\gamma = n - 1 = 21-1 = 20$

الجدولية $T_{20,0.05} = 1.72$

النتيجة:

استخدم اختبار T لفحص تساوي متوسطات الاختبار الأول والثاني؟

بما أن القيمة المحسوبة (0) > أقل من القيمة الجدولية (1.72) وهي تقع في منطقة القبول، إذاً نقبل الفرضية الصفرية، ويتبين انه ليس هناك فرقاً بين درجة الطالب في الاختبار الأول ودرجته في الاختبار الثاني.

مثال (5-7) أراد باحث مقارنة الرواتب التي يستلمها مجموعة من الموظفين من الأزواج

والزوجات وكانت كما يلي:

الأزواج	الزوج	الزوجة	Di=Stat1-Stat2	di ²
1	240	230	10	100
2	260	270	-10	100
3	250	240	10	100
4	230	230	0	0
5	280	270	10	100
6	220	230	-10	100
7	240	250	-10	100
8	260	250	10	100
9	240	230	10	100
10	250	220	30	900
المجموع			50	1700
d'			5	

$$S_d^2 = \frac{d_i^2 - (\underline{d'})^2}{n - 1} = \frac{1697.5}{9} = 188.6111$$

$$S_d = 13.73358$$

$$T = \frac{\bar{X} - \mu_0}{\frac{S_d}{\sqrt{n}}}$$

$$T = (5-0) / (13.73/3.16) = 5/4.43 = 1.13 \text{ المحسوبة}$$

$$\gamma = n - 1 = 10 - 1 = 9$$

$$T_{9,0.05} = 1.83 \text{ الجدولية}$$

النتيجة:

استخدم اختبار T لفحص تساوي متوسطات دخل الأزواج والزوجات ؟
بما أن القيمة المحسوبة (1.13) > أقل من القيمة الجدولية (1.83) وهي تقع في منطقة القبول، إذاً نقبل الفرضية الصفرية، ويتبين أنه ليس هناك فرقاً بين دخل الأزواج والزوجات.

أهم الأساليب الإحصائية الشائعة واستخداماتها:

المقاييس البارامترية	المقاييس البارامترية	المقاييس البارامترية	المقاييس البارامترية
<ul style="list-style-type: none"> - يستخدم لتقدير ما إذا كان توزيعان تكراريان يختلف عن بعضهما بشكل دال. 	كاي	<ul style="list-style-type: none"> - تحديد ما إذا كان متوسطان أو نسبتيان، أو معاملا ارتباط مختلفان عن بعضهما. - تحديد ما إذا كان متوسط واحد أو نسبة واحدة أو معامل ارتباط واحد يختلف عن تلك العلاقة للمجتمع. 	<p>الختبار (z) Z-test</p> <p>الختبار (ت) t-test</p>
<ul style="list-style-type: none"> - يستخدم لقياس ما إذا كان متوسطين غير مرتبطين يختلفان بشكل دال 	Mann-Whitney U test	<ul style="list-style-type: none"> - تحديد ما إذا كانت درجات المتوسط في عنصر أو أكثر تختلف عن بعضها. - ما إذا كان هناك تفاعل دال بين العناصر المختلفة. - يقيس إذا ما إذا كانت التباينات Variances مختلفة عن بعضها. 	<p>تحليل التباين Analysis of variance</p> <p>One way Anova</p> <p>Two way Anova</p>

اختبارات تستخدم بعد تحليل التباين. Duncan's Scheffe's Tukey	تستخدم إذا ظهرت قيمة F دالة وذلك بهدف اختبار الدلالة الإحصائية للفروق بين متوسطات مجموعات محددة	Wilcoxon signed test	يستخدم لقياس ما إذا كان متوسطين مترتبين مختلفان بشكل ذات
تحليل التباين Analysis of Covariance (Anacova)	مشابه في الاستخدام لأسلوب تحليل التباين إلا أنه يمكن من ضبط متغير مستقل أو أكثر في المتغير التابع.	Kruskal-Wallis test	يستخدم لتقدير ما إذا كان 3 قيم أو أكثر للمتوسطات في عنصر واحد تختلف بدلالة إحصائية
Trend Analysis	لاختبار الاتجاه المفترض		
Confidence limits	يستخدم لتقدير قيمة في المجتمع بالاعتماد على القيمة المعروفة للعينة.		

الأساليب الإحصائية لحساب الفروق:

مستويات القياس (القياس هنا للمتغير التابع)			عدد المتغيرات المستقلة	
اسمي	رتبي	لفوي أو لسي		
كاي تربيع لحسن المطابقة	Smirnov	t-test للعينة الواحدة	1	عينة واحدة
كاي تربيع للارتباط للعينات المستقلة ؛ Fisher exact test	Man Whitney U-Test. Median Test اختبار الوسيط	t-test للعينات المستقلة.	1	عينتان مستقلتان
كاي تربيع لنسبتين بيانات غير مستقلة	Wilcoxon Signed Rank Test. Sign Test. الإشارة	t-test للعينات المستقلة. Randomization Test. Walsh Test.	1	عينتان غير مستقلتان
كاي للعينات المستقلة	Kruskal Walliss كروسكال واليس Median test	Analysis of Variance (One Way Anova). تحليل التباين أحادي الاتجاه.	1	أكثر من عينتين مستقلتين

Anova (one way) اختبارات تستخدم بعد تحليل التباين. Duncan's Multiple- range, Scheffe's , test., Tuky	Friedman Test اختبار فريدمان	Chochran O test Friedman Test (البيانات الثنائية).	1	أكثر من عينتين غير مستقلتين
Factor analysis التحليل العاملي 2 way Anova تحليل التباين ثنائي الاتجاه. : Ancova		كاي تربيع	2 أو أكثر	عينتان أو أكثر

3-5 استخدام برمجية SPSS في حل المسائل.

1- اختبار الفرضية المتعلقة بوسط حسابي واحد.

اختبار T للعينة الواحدة One Sample T-Test

إذا كانت العلامات أدناه هي علامات الطلاب في مادة قواعد البيانات استخدم اختبار T لفحص وجود فرق بين متوسط درجة اختبار مادة قواعد البيانات (15.9286) وبين المتوسط الطبيعي وهو (12.5).

يمكن صياغة سؤال الدراسة بأحد الأشكال التالية:

هل هناك فرق بين متوسط درجة الطلاب وبين المتوسط الطبيعي؟

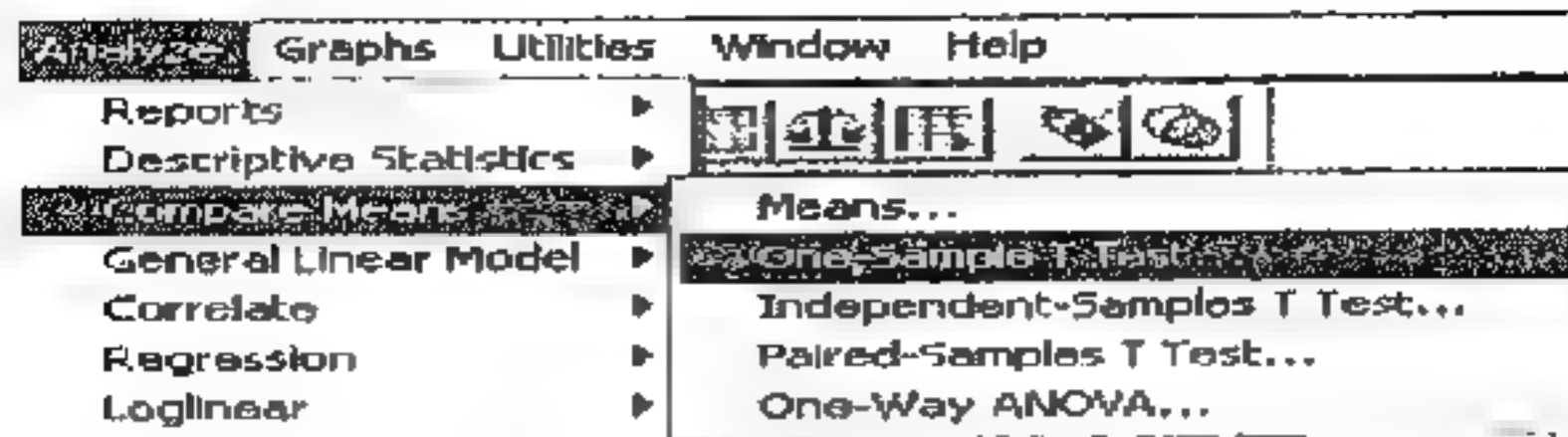
نرفض الفرضية إذا كانت دلالة قيمة (Sig.(2-tailed) أكبر من المستوى المقبول (0.05)

وهذا يعني أن المتوسط لا يساوي القيمة الثابتة a.

* ادخل البيانات كما هو مبين أدناه:

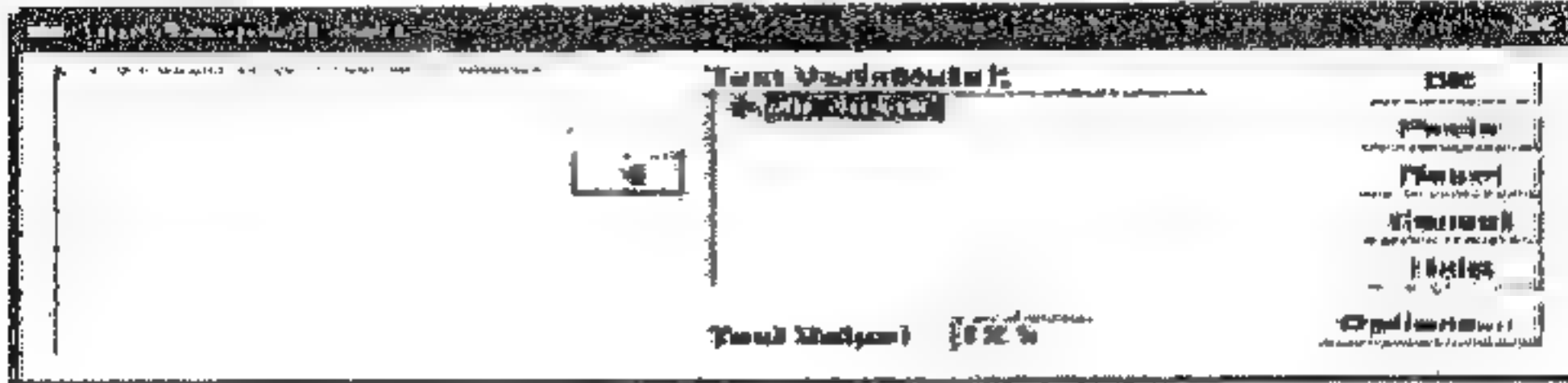
	db mark	var	var	var	var
1	14.00				
2	14.00				
3	20.00				
4	17.00				
5	17.00				
6	19.00				
7	20.00				
8	18.00				
9	14.00				
10	17.00				
11	8.00				
12	19.00				
13	14.00				
14	14.00				
15					

Analyze - Compare Means - One-Sample T Test...



ضع المتغير db_mark في خانة Test Variable(s):

وضع القيمة 12.5 في خانة Test Value: - ثم اضغط زر OK



تظهر النتيجة التالية:

T-Test

One-Sample Statistics

	عدد المشاهدات N	Mean الوسط الحسابي	Std. Deviation الانحراف المعياري	Std. Error Mean الخطأ المعياري
DB_MARK	14	15.9286	3.1736	.8482

One-Sample Test

	Test Value = 12.5					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
DB MARK	4.042	13	.001	3.4286	1.5962	5.2609

النتيجة:

تم حساب متوسط الفرق بين المتغير والقيمة المفترضة (Mean Difference) = 3.4286 والذي يشير إلى أن الوسط الحسابي لعينة الطلاب (15.9286) كان أعلى من المستوى الطبيعي (12.5)، وكان الفرق - 3.4286 وهل هذا الفرق كافٍ لأن نقرر أنه ذو دلالة إحصائية؟

$$H_0: \mu = 12.5$$

$$H_1: \mu \neq 12.5$$

عما أن مستوى المعنوية $\text{Sig. (2-tailed)} = 0.001$ وهي أقل من $>$ مستوى الدلالة

$\alpha = 0.05$ ، إذا تقبل الفرضية البديلة ونرفض الفرضية الصفرية

استخدم اختبار T لفحص وجود فرق بين متوسط درجة اختبار مادة قواعد البيانات وبين المتوسط الطبيعي (12.5)، وقد تبين من خلال النتائج الموضحة في الجدول أعلاه أن متوسط علامات الطلبة كان أعلى من المتوسط الطبيعي (12.5) فقد بلغ متوسط علامات الطلاب (15.93) بإحرف معياري (3.17) وقد بلغت قيمة $t = 4.042$ وهي ذات دلالة احصائية عند مستوى أقل من 0.05

كانت قيمة $t = 4.042$ وبلغ مستوى دلالتها $\text{Sig. (2-tailed)} = 0.001$ قيمة صغيرة جداً وهي أقل من المستوى المقبول لدينا ($\alpha = 0.05$) وهذا يعني أن متوسط الاختبار في مادة قواعد البيانات $\neq 12.5$ بل هو أعلى من ذلك.

القيمة المحسوبة $t = 4.042$

القيمة الحرجة $[-2.160 \quad +2.160]$

إن القيمة المحسوبة تقع في منطقة رفض الفرضية الصفرية يعني قبول البديلة.

المطرب: إذا كانت العلامات أدناه هي علامات الطلاب في مادة الإحصاء استخدم اختبار T لفحص وجود فرق بين متوسط درجة اختبار مادة الإحصاء (15.7619) وبين المتوسط الطبيعي وهو (12.5).

$$H_0: \mu = 12.5$$

$$H_1: \mu \neq 12.5$$

	stat1_m
1	20.00
2	20.00
3	18.00
4	15.00
5	16.00
6	12.00
7	16.00
8	18.00
9	18.00
10	18.00
11	14.00
12	16.00
13	12.00
14	17.00
15	14.00
16	19.00
17	11.00
18	10.00
19	19.00
20	12.00
21	17.00

Data View

النتيجة: بما أن مستوى المعنوية $\text{Sig. (2-tailed)} = 0.000$ ومستوى الدلالة $\alpha = 0.05$ ، إذا نقبل الفرضية البديلة ونرفض الفرضية الصفرية يعني أن متوسط الاختبار في مادة الإحصاء $\neq 12.5$ بل هو أعلى من ذلك.

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
STAT_MAR	21	15.7619	3.0480	.6651

One-Sample Test

	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
STAT_MAR	4.904	20	.000	3.2619	1.8745	4.6494

2- اختبار الفرضيات حول الفرق بين وسطين حسابيين (للبينات المترابطة)

اختبار T للعينات المترابطة Paired Sample T-Test

السؤال هو: هل تتساوى متوسطات الطلاب بالاختبارين؟

نرفض الفرضية إذا كانت دلالة قيمة Sig. (2-tailed) أقل من المستوى المقبول (0.05)

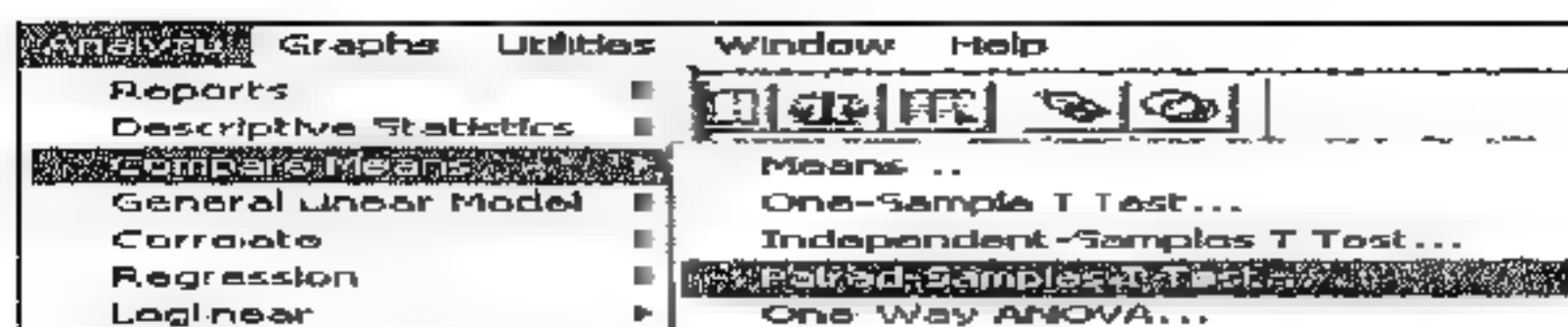
وهذا يعني أن المتوسطين غير متساويين.

* ادخل البيانات كما هو مبين أدناه:

	stat1_m	stat2_m
1	20.00	15.00
2	20.00	14.00
3	18.00	20.00
4	15.00	17.00
5	16.00	17.00
6	12.00	18.00
7	15.00	20.00
8	18.00	18.00
9	18.00	14.00
10	18.00	17.00
11	14.00	8.00
12	16.00	18.00
13	12.00	14.00
14	17.00	14.00
15	14.00	12.00
16	19.00	20.00
17	11.00	10.00
18	10.00	10.00
19	19.00	20.00
20	12.00	16.00
21	17.00	20.00

Data View Variable View

Analyze - Compare Means - Paired-Samples T Test...



ضع المتغيرات stat1_m و stat2_m في خانة Paired Variables:

ثم اضغط زر OK



تظهر لدينا المخرجات التالية:

المطلوب: إذا كانت العلامات أدناه هي علامات الطلاب في مادة الإحصاء في الاختبار الأول والاختبار الثاني لفحص هل لهما نفس المتوسط الحسابي، استخدم اختبار لفحص وجود فرق بين متوسطي درجات الاختبارين في مادة الإحصاء.

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

T-Test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	STAT1_M	15.7619	21	3.0480	.6651
	STAT2_M	15.7619	21	3.6319	.7925

Paired Samples Correlations

	N	Correlation	Sig.
Pair 1 STAT1_M & STAT2_M	21	.505	.020

Paired Samples Test

		Paired Samples Statistics					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	STAT1 M. STAT2 M	.0000	3.3815	.7335	-1.5302	1.5302	.000	20	1.000

النتيجة:

استخدم اختبار T لفحص تساوي متوسطات الاختبار الأول والثاني؟
 بما أن مستوى المعنوية $\text{Sig. (2-tailed)} = 1.000$ ومستوى الدلالة $\alpha = 0.05$ ، إذا نرفض الفرضية البديلة ونقبل الفرضية الصفرية.
 تبين أنه ليس هناك فرقاً بين درجة الطالب في الاختبار الأول ودرجته في الاختبار الثاني فقد بلغت قيمة $t = (0.000)$ وهي غير دالة إحصائياً على مستوى $\alpha = 0.05$ حيث حسب البرنامج متوسط الفرق بين درجة الطالب في الاختبار الأول ودرجته في الاختبار الثاني والذي بلغ (0.0000) .

3- اختبار الفرضيات حول السفرق بين وسطين حسابيين (للبينات المستقلة)

اختبار T للعينات المستقلة Independent-Sample T-Test

فحص فرضية متعلقة بمساواة متوسط متغير ما لعينتين مستقلتين، وله شكلان الأول في حالة افتراض تساوي تباين العينتين، والثاني في حالة افتراض أن تباين العينتين غير متساوي.
 إذا كانت علامات الطلاب والطالبات في مادة الإحصاء في شعبة ما كما هو مبين أدناه، فهل يختلف تحصيل الذكور عن الإناث في هذه الشعبة؟
 هل يرتبط تحصيل الطلبة حسب الجنس؟
 الفرضية المطلوب اختبارها:

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

للاستخدام الاختبار يكون هناك متغير التجميع (Grouping Variable) وهو المتغير الذي يقسم العينة الكلية إلى عينتين جزئيتين غير متداخلتين، مثل متغير الجنس الذي يقسم العينة

الكسبة إلى عينة ذكور وعينة إناث، والمتغير الثاني يسمى متغير الاختبار (Test Variable) أو المتغير التابع، وهو متغير كمي .

والهدف من هذا الاختبار هو معرفة ما إذا كان متوسط متغير الاختبار لفئة متغير التجميع الأولى (الذكور) مساوية لمتوسط متغير الاختبار لدى الفئة الثانية (الإناث) من متغير التجميع.

السؤال هو: هل يختلف تحصيل الطلاب في مادة الإحصاء باختلاف جنسهم؟

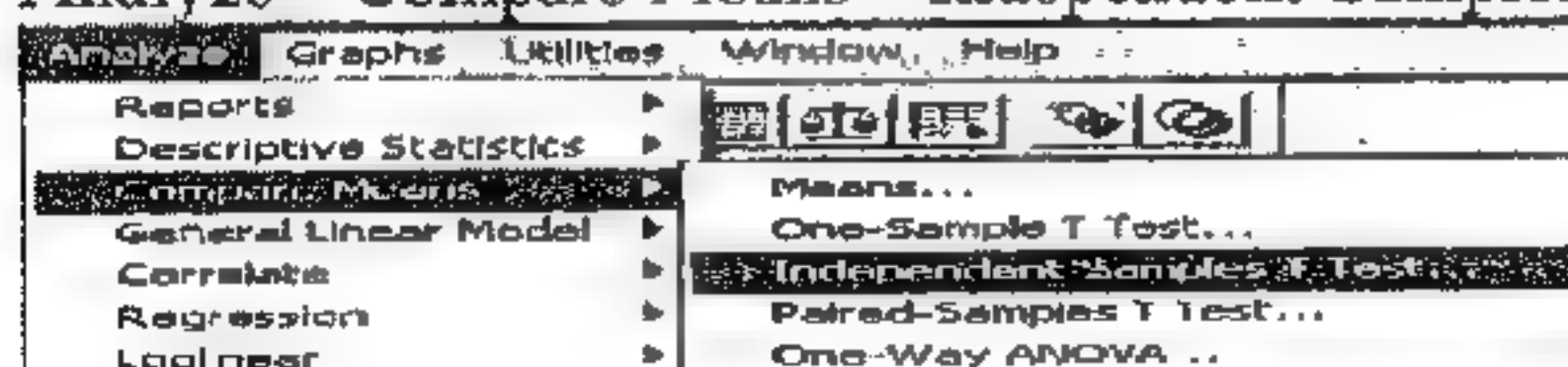
أو هل يرتبط تحصيل الطلاب في مادة الإحصاء بالجنس؟

نرفض الفرضية الصفرية إذا كانت دلالة قيمة مستوى الدلالة (Sig.(2-tailed) أقل من المستوى المقبول (0.05) وهذا يعني أن المتوسطين غير متساويين، وذلك بعد تحديد قيمة t المستخدمة بناء على نتيجة اختبار Levene Test لمساواة تباين العينتين.

* ادخل البيانات كما هو مبين أدناه:

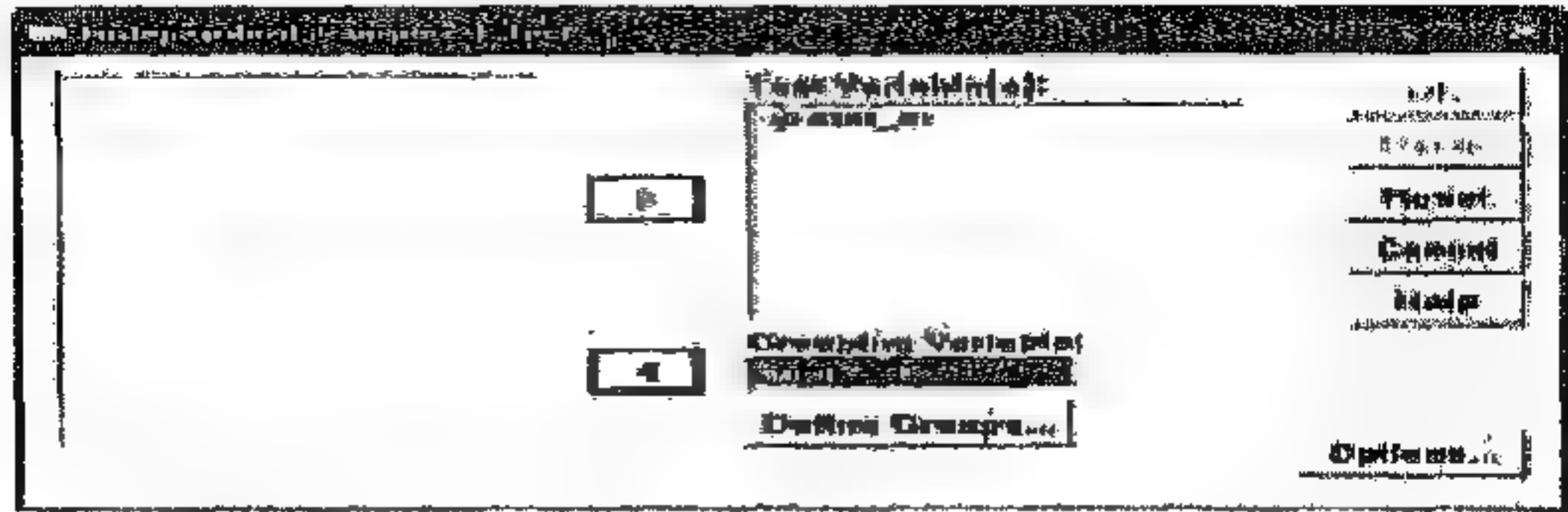
	sex	stat_m
1	1.00	16.00
2	1.00	16.00
3	1.00	16.00
4	1.00	16.00
5	1.00	12.00
6	1.00	16.00
7	1.00	16.00
8	1.00	16.00
9	1.00	16.00
10	1.00	14.00
11	2.00	16.00
12	2.00	12.00
13	2.00	17.00
14	2.00	14.00
15	2.00	19.00
16	2.00	11.00
17	2.00	10.00
18	2.00	19.00
19	2.00	12.00
20	2.00	17.00
21		

Analyze - Compare Means - Independent-Samples T Test...

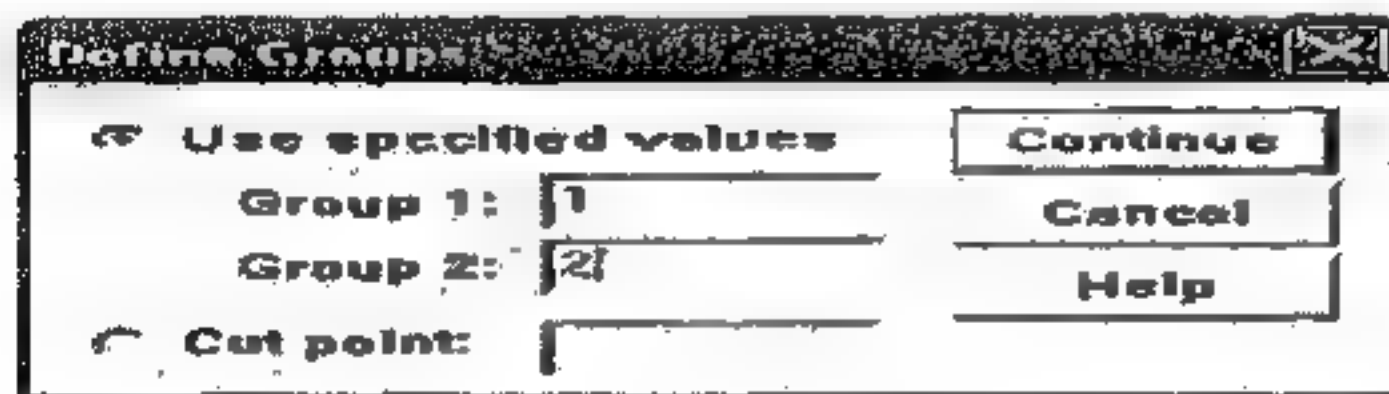


Test Variable(s): وضع المتغير stat_m في خانة

Grouping Variable: وضع المتغير sex في خانة



انقر زر Define Groups، ضع القيمة 1 في Group1، وضع القيمة 2 في Group2،
، ثم اضغط زر Continue، ثم اضغط زر Ok



تظهر لدينا المخرجات أدناه:

الإحصاءات الوصفية لكل عينة

Group Statistics

	SEX	N	Mean	Std. Deviation	Std. Error Mean
STAT_M	1.00	10	16.0000	2.0548	.6498
	2.00	10	14.7000	3.3350	1.0546

Independent Sample t Test

		Levene's Test for Equality of Variances		t Test for Equality of Means						
		f	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
STAT_M	Equal variances assumed	5.356	.033	1.049	18	.308	1.3000	1.2387	-1.3025	3.9025
	Equal variances not assumed			1.049	14.982	.311	1.3000	1.2387	-1.3407	3.9407

النتيجة:

تم اختبار تجانس التباين للفتتين بواسطة اختبار Levene Test ، حيث حُسبت قيمة F وكانت = 5.356 ومستوى دلالتها فكانت = 0.033 وذلك لتحديد أي الاختبارين سنستخدم، هل سنستخدم اختبار T في حالة تساوي التباين Equal variances assumed أم اختبار T في حالة عدم تساوي التباين Equal variances not assumed

استخدم اختبار T لفحص السؤال "هل يختلف تحصيل الطلاب عن الطالبات؟" أو "هل يرتبط تحصيل الطلاب بالجنس؟"

وقد وجد من خلال نتائج هذا الاختبار أنه ليس هناك فرقاً بين تحصيل الطلاب والطالبات حيث بلغت قيمة $t=1.049$ وهي ليست ذات دلالة إحصائية على مستوى أقل من 0.05

نختار قسيمة t ومستوى دلالتها بناء على اختبار F لنقرر هل نختار اختبار T في حالة افتراض تساوي التباين أم اختبار T في حالة افتراض عدم تساوي التباين.

في هذه الحالة نختار اختبار T في حالة افتراض عدم تساوي التباين لأن مستوى دلالة $\alpha=0.033$ وقيمة $F=5.356$ أقل من 0.05 وبالتالي فإن تباين الفئتين غير متساوي.

وقد بلغ متوسط تحصيل الطلاب 16 بانحراف معياري 2.0548 في حين بلغ متوسط تحصيل الطالبات 14.7 بانحراف معياري 3.3350 حيث يتبين أن لا فرق في التحصيل بسبب الجنس.

ملاحظة: استخدام نقطة القطع Cut Point، والرسومات البيانية.

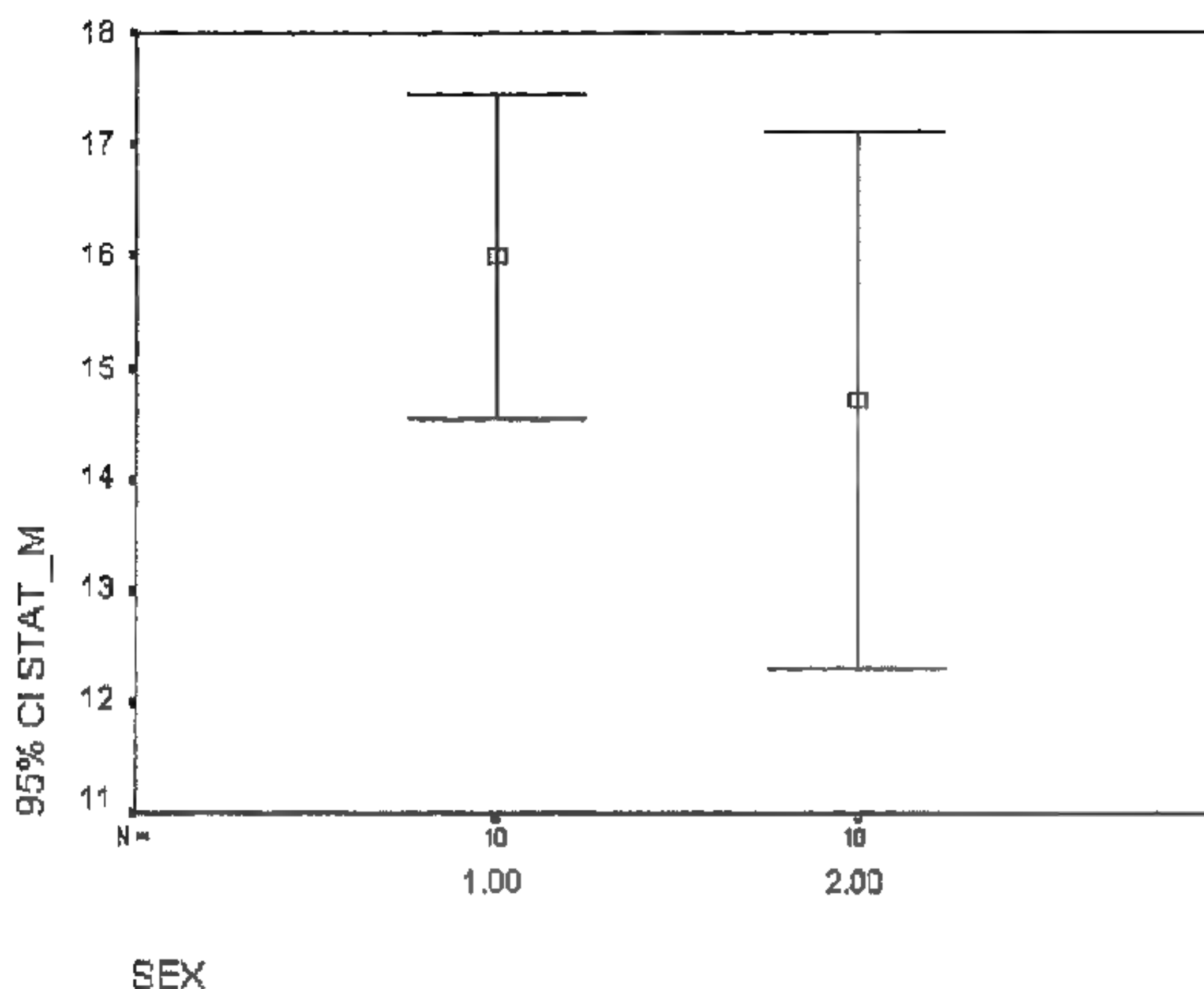
T-Test

Group Statistics

AGE	N	Mean	Std. Deviation	Std. Error Mean
STAT_M ≥ 20.00	8	14.8000	3.11677	1.10165
< 20.00	12	15.9167	2.90303	.72288

Independent Samples Test

	Levene's Test for Equality of Variances		Test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
STAT_M Equal variances assumed	.830	.374	-1.125	18	.275	-1.4167	1.25584	-4.08140	1.22805
Equal variances not assumed			-1.075	12.807	.302	-1.4167	1.31772	-4.28778	1.43443



4-5 تمارين Exercise

ضع رمز الإجابة الصحيحة في المربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										
الرقم	11	12	13	14	15	16	17	18	19	20
الإجابة										

1- علاقة Z ب σ_x كعلاقة T ب

أ- σ ب- σ^2 ج- S^2_x د- S_x'

2- أي من التالي يمكن اعتباره فرضية إحصائية؟

أ- $\sigma = 10$ ب- $X = 27$ ج- $r = 0.2$ د- $S_x = 8$

- 3- أي من التالي يمكن اعتباره فرضية إحصائية؟
 أ- $\mu = 70$ ب- $X = 27$ ج- $r = 0.2$ د- $Sx = 8$
- 4- كم يجب أن يكون حجم العينة حتى يكون الخطأ المعياري للوسط الحسابي $\sigma_x = 10\%$ من الانحراف المعياري σ في المجتمع؟
 أ- $\sqrt{10}$ ب- 1000 ج- 10 د- 100
- 5- إذا كانت $Z=2.1$ وكانت الفرضية البديلة غير متجهة، فإنه بالإمكان :
 أ- رفض الفرضية الصفرية على مستوى $\alpha = 0.01$
 ب- رفض الفرضية الصفرية على مستوى $\alpha = 0.05$
 ج- رفض الفرضية الصفرية على مستوى $\alpha = 0.05$
 د- رفض الفرضية الصفرية على المستويين $\alpha = 0.01$ ، $\alpha = 0.05$
- 6- إذا كانت فترة الثقة 95% للوسط الحسابي تمتد ما بين 53.6 - 76.4 فأى من الفرضيات الإحصائية التالية تعتبر مرفوضة على مستوى $\alpha = 0.05$ ، مع الأخذ بالحسبان أن فرضيتها البديلة غير متجهة:
 أ- $\mu = 51.6$ ب- $\mu = 65$ ج- $\mu = 75.4$ د- $\mu = 76.1$
- 7- افترض أن الفرضية الصفرية التالية $\mu = 60$ كانت مرفوضة على مستوى $\alpha = 0.01$ فهل هذا يعني أن القيمة 60 تقع:
 أ- ضمن فترة الثقة 95%
 ب- ضمن فترة الثقة 99%
 ج- خارج فترة الثقة 95%
 د- خارج فترة الثقة 99% ، 95%
- 8- يستخدم اختبار T كاختبار للفرضية الصفرية حول الوسط الحسابي إذا كانت غير معلومة.
 أ- n ب- \bar{x} ج- σ د- α
- 9- أيهما أكثر شيوعاً في الاستخدام لفحص الفرضية حول الوسط الحسابي؟
 أ- اختبار Z ب- اختبار T ج- اختبار F د- اختبار χ^2

10- تكون قيم T الحرجة عند نفس مستوى الدلالة:

أ- أكبر للاختبار بذييل. ب- أكبر للاختبار بذييلين.

ج- نفس القيمة. د- لا شيء مما ذكر.

11- ما هي قيم T الحرجة في ذيل واحد وفي ذيلين إذا كانت $n=21$ ، $\alpha = 0.05$

أ- (1.721, 2.080) ب- (2.528, 2.845)

ج- (1.725, 2.086) د- (2.518, 2.0831)

12- ما هي قيم T الحرجة في ذيل واحد وفي ذيلين إذا كانت $n=30$ ، $\alpha = 0.01$

أ- (2.457, 2.760) ب- (1.699, 2.045)

ج- (2.462, 2.756) د- (1.697, 2.042)

13- درست إحدى شعب الإحصاء المتقدم التي عدد أفرادها 45 طالباً وحدة اختبار

الفرضيات حول الوسط الحسابي باستخدام مختبر الحاسوب، وعند إجراء اختبار يقيس

مهارتهم في هذه الوحدة وجد أن متوسط أدائهم = 65.3 فإذا كانت النتائج المعينة لجميع

الطلاب الآخرين هي 50 للوسط الحسابي و 12 للانحراف المعياري، فهل هناك ما يدعو

إلى القول بأن نتائج الطلبة الذين استخدموا المختبر أعلى من نتائج الآخرين؟ استخدم

$$\alpha = 0.01$$

أ- نعم لأن $Z = 8.56$ وهي تقع في منطقة رفض الفرضية الصفرية.

ب- نعم لأن $Z = 8.56$ وهي تقع في منطقة قبول الفرضية الصفرية.

ج- لا لأن $Z = 8.56$ وهي تقع في منطقة رفض الفرضية البديلة.

د- لا لأن $Z = 8.56$ وهي تقع في منطقة قبول الفرضية البديلة.

14- فحص فادي فرضيته الصفرية $H_0: \mu = 10$ بالاستعانة باختبار بذييلين $\alpha = 0.05$

مستخدماً عينة حجمها 25 ، وفحص جمال الفرضية نفسها بالاستعانة باختبار بذييلين $\alpha =$

0.05 مستخدماً عينة حجمها 100 فإن احتمال ارتكاب الخطأ من النوع الأول يكون

أكبر ل:

أ- فادي ب- جمال ج- لا يمكن حسابه د- نفس الاحتمال.

15- فحص فادي فرضيته الصفرية $H_0: \mu = 10$ بالاستعانة باختبار بديين $\alpha = 0.05$ مستخدماً عينة حجمها 25 ، وفحص جمال الفرضية نفسها بالاستعانة باختبار بديين $\alpha = 0.05$ مستخدماً عينة حجمها 100 ، إذا كانت القيمة الحقيقية $\mu = 12$ فأيهما يكون احتمال ارتكابه للخطأ من النوع الثاني أكبر؟
أ- فادي ب- جمال ج- لا يمكن حسابه د- نفس الاحتمال.

16- فحص فادي فرضيته الصفرية $H_0: \mu = 10$ بالاستعانة باختبار بديين $\alpha = 0.05$ مستخدماً عينة حجمها 25 ، وفحص جمال الفرضية نفسها بالاستعانة باختبار بديين $\alpha = 0.05$ مستخدماً عينة حجمها 100 ، إذا كانت القيمة الحقيقية $\mu = 12$ فأيهما يكون اختبار أقوى؟
أ- فادي ب- جمال ج- لا يمكن حسابه د- نفس الاحتمال.

17- يدعي عميد الكلية بأن توجيه إنذارات إلى الطلاب يقلل من غياب الطلاب في الجامعة. فإذا علم أن متوسط غياب الطلاب في الكلية في الأعوام المنصرمة كان 4.6 أيام للطلاب الواحد. وبعد أن استخدم هذا الأسلوب لهذا العام مع صف عدد طلابه 25 طالباً وجد أن متوسط الغياب كان 3.7 أيام للطلاب الواحد بانحراف معياري غير متحيز $= 1.2$ يوم. هل تدعم هذه البيانات صحة ادعاء العميد؟ استخدم $\alpha = 0.05$
أ- نعم، لأن $T = -3.75$ وهي تقع في منطقة رفض الفرضية الصفرية.
ب- لا، لأن $T = -3.75$ وهي تقع في منطقة رفض الفرضية الصفرية.
ج- نعم، لأن $T = -3.75$ وهي تقع في منطقة رفض الفرضية البديلة.
د- لا، لأن $T = -3.75$ وهي تقع في منطقة رفض الفرضية البديلة.

18- فيما يلي علامات عشرة طلاب على مقياس الاتجاهات، جرى تطبيقه قبل وبعد مشاهدة فيلم عن العنف المدرسي:

اسم الطالب	قبل	بعد
فادي	10	14
جمال	17	31
نبيل	12	17
فايز	19	17
جمعه	10	13
محمد	11	24
احمد	20	25
صلاح	13	14
سمير	17	16
يوسف	20	15

هل تحسنت اتجاهاتهم نحو القضية بعد مشاهدة الفيلم؟ $\alpha = 0.05$

أ- لا وذلك لأن $T=1.92$ وهي ذات دلالة إحصائية على مستوى $\alpha = 0.05$ للاختبار
بذيل واحد.

ب- نعم وذلك لأن $T=1.92$ وهي ذات دلالة إحصائية على مستوى $\alpha = 0.05$
للاختبار بذيل واحد.

ج- لا وذلك لأن $T=1.92$ وهي ذات دلالة إحصائية على مستوى $\alpha = 0.05$
للاختبار بذيلين.

د- نعم وذلك لأن $T=1.92$ وهي ذات دلالة إحصائية على مستوى $\alpha = 0.05$
للاختبار بذيلين.

19- درست مجموعتان مادة الإحصاء المتقدم بطريقتين مختلفتين (نظري، عملي) وعند انتهاء
فترة التدريس أجري اختبار تحصيلي للمجموعتين والجدول أدناه يبين النتائج:

الانحراف المعياري غير المتحيز S_x	الوسط الحسابي \bar{X}	عدد الأفراد n	
8	65	30	طريقة النظري
9	55	25	طريقة عملي

هل تختلف الطريقتان في تأثيرهما على التحصيل؟ علماً بأن المجموعتين قد اختيرتا بطريقة عشوائية.

- أ- نعم وذلك لأن $T=4.361$ وهي ليست ذات دلالة إحصائية على مستوى $p<0.001$
 ب- لا وذلك لأن $T=4.361$ وهي ذات دلالة إحصائية على مستوى $p<0.001$.
 ج- نعم وذلك لأن $T=4.361$ وهي ذات دلالة إحصائية على مستوى $p<0.001$.
 د- لا وذلك لأن $T=4.361$ وهي ليست ذات دلالة إحصائية على مستوى $p<0.001$

1- سحبت عينتين عشوائيتين من مجتمعين طبيعيين حجم الأولى 5 والثانية 4 والبيانات موضحة في الجدول التالي اعتبر أن كانت توجد فروق جوهرية بين متوسط مجتمعيهما بمستوى دلالة $\alpha=0.05$.

العينة الأولى	العينة الثانية
3.1	2.3
4.4	1.4
1.2	3.7
1.7	8.9
3.4	

2- إحدى شركات المنتجات النفطية أنتجت نوع معين مطور من مادة الكاسولين المحسن لزيادة عدد الكيلو مترات من المسافات المقطوعة. ولاختبار هذا، سحبت عينة عشوائية من 10 سيارات وسارت باستخدام الكاسولين المحسن والعادي والجدول التالي يبين المسافات المقطوعة من قبل العشر سيارات. اعتبر على مستوى 0.05 أنه لا توجد فروق جوهرية وإن كانت هناك فروق قدرها بحدود الثقة السابقة في السؤال.

كاسولين	
عادي	محسن
24.9	25.7
18.8	20.5
27.7	28.4
13.0	3.7
17.8	18.8
11.3	12.5
27.6	28.4
8.2	8.1
23.1	23.1
9.9	10.4

3- الجدول التالي يمثل بيانات عن الاجور التي تتقاضاها عينتين عشوائيتين و المسحوبة من مجتمعين طبيعيين أحدهم المشتغلين لدى الدولة والثانية لدى القطاع الخاص.

الاجور لدى القطاع الخاص	الاجور لدى قطاع الدولة	
35	30	حجم العينة
35558.97	3333520	الوسط الحسابي
14940.88	15129.09	النحراف المعياري

احسب فترة ثقة 95% للفرق بين متوسطي مجتمعهما.

4- اجري اختبار في احدى المسابقات ومع شعبتين مختلفتين الاولى مكونة من 40 طالباً والثانية 30 طالباً وجد أن متوسط علامات الطلبة في الشعبة الاولى 65 درجة بانحراف معياري قدره 10 درجة ومتوسط علامات الطلبة في الشعبة الثانية 57 درجة بانحراف معياري 6 درجة اختبر إن كان هناك فروق جوهرية بين مستوى الطلبة في الشعبتين على مستوى معنوية $\alpha = 0.05$.

5- إذا كان متوسط الزيادة في وزن 12 فأرة بعد تغذيتها بطريقة معينة لمدة معينة هو 145 غم وبانحراف قياس للوسط الحسابي مقداره 2.3 غم وبمستوى احتمال 5% هل يمكن القول أن متوسط الزيادة في الوزن نتيجة التغذية على الطريقة لا تقل عن 150 غم اختبر ذلك.

6- شركة توزيع المحروقات ارادت تقدير الوسط الحسابي للمجتمع بمحدود ثقة 95% سحبت عينة عشوائية مكونة من 100 عائلة مستهلكة للسولار كان معدل استهلاكها ما يعادل 1103 غالون بانحراف معياري 327.8. احسب الوسط الحسابي للمجتمع الذي سحبت منه العينة ثم فسر إذا استهلكت عائلة منفردة 800 غالون هل يمكن اعتبار ذلك ممكن اختبر ذلك.

7- إحدى شركات الاتصالات اجرت بحثا حول المكالمات الطويلة فوجد ان معدل ما يدفعه المواطن للمكالمة الطويلة يساوي 17.10 دولار في الشهر. وبانحراف معياري 9.80 دولار. سحبت عينة عشوائية لـ 50 قائمة تلفون.

- أوجد احتمال أن تكون مدة المكالمة اكبر من 20 دولار.

- اختبر أن $H_0: \mu = 21$ ضد $H_1: \mu > 21$ مستخدماً $\alpha = 0.05$

8- أجريت مقارنة للأسعار في مدينتين مثل اليابان وأمريكا فإذا كان سعر الفرق لبعض المواد التجارية في كل من الدولتين موضحة في الجدول أدناه:

	أمريكا	اليابان
حجم العينة	$N_1=50$	$n_2=30$
معدل سعر	1154.5	1224.3
الانحراف المعياري	1989	1843

اختبر أن فروقات الاسعار هي أكبر من 200 دولار بمستوى دلالة $\alpha = 0.05$

9- احسب التقدير النقطي لعينة عشوائية حجمها 12 مسحوبة من مجتمع μ

المجموعة الثانية	المجموعة الأولى	
35015	38.75	الوسط الحسابي
2.7	3.2	الانحراف المعياري
100	100	حجم العينة

وكانت أفراد العينة هي:

X: 9 6 5 3 4 7 8 9 10 3 12 6

10- في إحدى المدارس الأساسية سحبت عينة عشوائية من الطلبة الذين سيعملون النظارات الطبية حجمها 20 طالباً فوجد أن 6 منهم يستخدم النظارات الطبية. فما تقديرك للسته الذين يستعملون النظارات الطبية في تلك المدرسة.

11- سحبت عينة عشوائية حجمها 400 مفردة من مجتمع انحراف القياسي 30 ومعدل 160 احسب فترة ثقة 95% للوسط الحسابي للمجتمع الذي سحبت منه هذه العينة.

12- سحبت عينة عشوائية من إحدى المصانع الكهربائية حجمها 25 مصباحاً فكان الوسط الحسابي لأعمار هذه المصابيح 890 ساعة. احسب فترة ثقة 90% لمعدل أعمار المصابيح المنتجة في هذا المصنع على أن الانحراف المعياري لإنتاجية هذا المصنع هو 35 ساعة.

13- سحبت عينة عشوائية في إحدى مصانع الخيوط حجمها 60 خيطاً فوجد أن معدل قوة هذه الخيوط 90.4 كغم بانحراف معياري 7.5 كغم. أوجد فترة ثقة 98% لمعدل قوة جميع الخيوط التي ينتجها ذلك المصنع.

14- سحبت عينة عشوائية مكونة من 25 طالباً من جامعة مؤته لتقدير معدل المصروف الشهري لهم. فوجد أن مصروفهم الشهري بالدينار الأردني كما يلي:

38 51 49 38 36 35 44 50

43 41 44 38 33 45 50 51

30 50 45 40 30 44 39 49 52

- ما هو تقديرك لمعدل المصروف الشهري لجميع طلبة جامعة مؤته.

- احسب فترة ثقة 95% لمعدل المصروف إذا كان المصروف يخضع لتوزيع طبيعي.

15- سجلت قياسات الحموضة (PH) لعينات من ماء المطر في 10 مواقع في منطقة صناعية فكانت 3.9 3.1 5.1 3.8 4.5 3.2 4.8 3.9 4.1 3.6 احسب فترة ثقة 95% لمعدل حموضة ماء المطر لكل المناطق.

16- إذا كانت أجور مندوبي المبيعات لكل من الذكور والإناث تخضع لتوزيع طبيعي متباين 100 للذكور و144 للإناث علما بأن المجتمعين للذكور والإناث مستقلين عن بعضهما. سحبت عينة عشوائية حجمها 170 وانحراف معياري 9 وعينة عشوائية ثانية (إناث) حجمها 24 بوسط 162 وانحراف معياري 10.
- أوجد فترة ثقة 98% لكل من μ_1, μ_2
- احسب فترة ثقة للفرق.

17- الجدول التالي يبين التحصيل العلمي لمجموعة من الطلبة في مدينتين مختلفتين.

المدينة	حجم العينة	الوسط الحسابي	الانحراف المعياري
أ	200	79	11
ب	150	73	12

احسب فترة ثقة 95% للفرق بين معدلي تحصيل الطلبة.

18- الجدول التالي يمثل الأجور التي يتقاضاها عنتين عشوائيتين والمسحوبة من مجتمعين إحداهما المشتغلين لدى القطاع الخاص والثانية لدى الدولة.

	العينة الأولى (الخاص)	الثانية (الدولة)
حجم العينة	35	30
معدل الأجر بالدولار	35558.79	33335.20
الانحراف المعياري	14940.88	15129.09

احسب فترة ثقة 95% للفرق بين متوسطي مجتمعتهما.

19 البيانات التالية تمثل فروقات قياس ضغط الدم قبل وبعد تناول الدواء، فروقات ضغط الدم تساوي = 13 - 2 - 1 - 6 4 احسب الوسط الحسابي لمجتمع الفروقات بمعدل ثقة 90%.

20- تم اختيار مجموعتين من الطلبة في مادة الرياضيات وسحبت عينتين عشوائيتين من طلاب وطالبات إحدى المدارس وكانت النتائج كما يلي:

طالبات	طلاب	
85	18	الوسط الحسابي
4	5	الانحراف المعياري
12	10	حجم العينة

على افتراض أن المجتمعين يتوزعان قريبا من التوزيع الطبيعي وأن تباينهما غير معلوم ولكنهما متساويان احسب فترة ثقة 97% للفرق بين متوسط المجتمعين.

الفصل السادس

اختبار الفرضيات حول التباينات

Testing Hypothesis Inference Regarding Variances

- 1-6 اختبار فرضية تتعلق بالتباين لمجتمع واحد.
2-6 اختبار فرضية تتعلق بتساوي التباين لمجمعين مستقلين.
3-6 اختبار فرضية تتعلق بتساوي التباين لمجمعين غير مستقلين.
4-6 تمارين Exercise.

الفصل السادس

اختبار الفرضيات حول التباينات

Testing Hypothesis Inference Regarding Variances

1-6 اختبار فرضية تتعلق بالتباين لمجتمع واحد. Testing Hypothesis Inference Regarding Variances

اختبار الفرضيات حول التباين ضروري في الاختبارات الإحصائية التي تتطلب توفر

تجانس التباين.

- توزيع كاي تربيع (χ^2)

■ لا يأخذ قيماً سالبة.

■ القيمة العظمى للاحتمال فيه (68% من الحالات) تنحصر بين 0 - 1.

■ ملتو التواء موجب، وإن الالتواء يقل بزيادة درجات الحرية.

■ ليس له شكل محدد ولكن يعتمد شكله على درجات الحرية.

$$\chi^2 = \sum \frac{(Xr - \mu)^2}{\sigma^2}$$

X : الملاحظة على المتغير التابع.

μ : وسط المجتمع الإحصائي على المتغير التابع.

σ^2 : التباين في الملاحظات للمجتمع على المتغير التابع.

مثال (1-6) استخراج قيمة χ^2 التي تحصر 0.05 من المساحة في الذيل الموجب عندما

تكون درجات الحرية = 19

الحل: من جدول χ^2 نبحث عن تقاطع 0.05 من المساحة مع درجات الحرية 19

فنجد قيمة $\chi^2 = 30.14$

استخدام توزيع χ^2 في اختبار الفرضيات حول التباين

$$\chi^2_{n-1} = \frac{(n-1)S^2}{\sigma^2}$$

S: التقدير غير المتحيز للانحراف المعياري n : عدد المشاهدات.

σ : الانحراف المعياري للمجتمع الاحصائي. $\gamma=n-1$

مثال (6-2) إذا كان الانحراف المعياري في أطوال الرجال في مجتمع معين الذين تتراوح أعمارهم بين 20-25 سنة يساوي 2.5 إنش، ويراد معرفة فيما لو كان هذا الانحراف المعياري هو نفسه للنساء من نفس العمر، فتم اختيار 41 امرأة من تتراوح أعمارهن بين 20-25 سنة وجرى قياس أطولهن فكان الانحراف المعياري لهن 2.133 إنش. استخدم هذه البيانات في فحص ادعاء الباحث بأن التباين في أطوال النساء يقل عما هو الحال عند الرجال من نفس العمر استخدم مستوى الدلالة $\alpha = 0.01$



الحل:

$$H_0: \sigma^2 = (2.5)^2 = 6.25$$

$$H_1: \sigma^2 < 6.25$$

بما أن مستوى الدلالة $\alpha = 0.01$ نبحث في جدول χ^2 تحت عمود 0.99 ودرجات

حرية 41-1=40 نجد أن القيمة الحرجة = 22.16

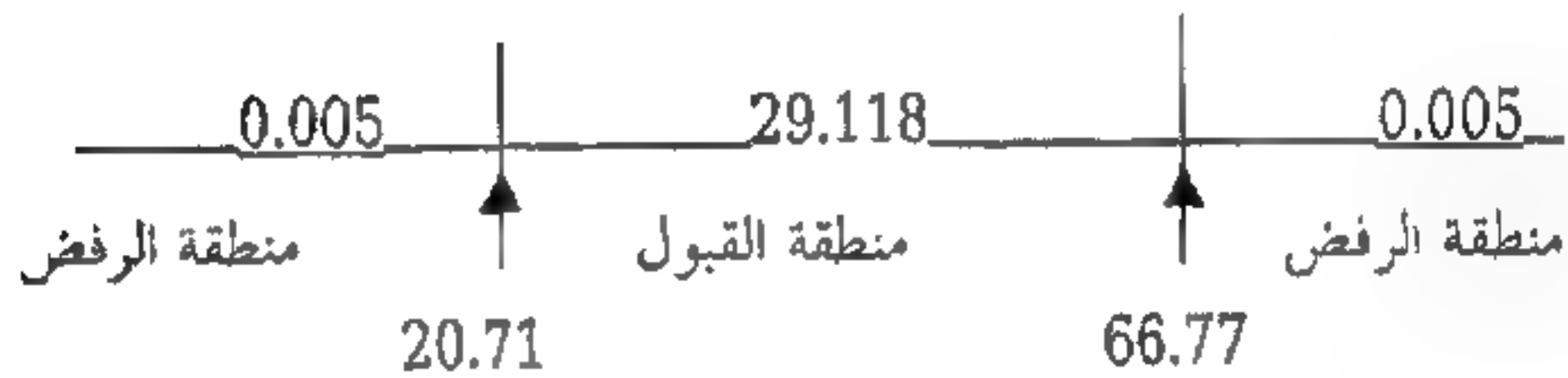
القيمة المحسوبة

$$\chi^2_{n-1} = \frac{(n-1)S^2}{\sigma^2} = \frac{(41-1)(2.133)^2}{6.25} = 29.118$$

النتيجة:

بما ان القيمة المحسوبة 29.118 أكبر من القيمة الحرجة = 22.16 لا نستطيع رفض الفرضية الصفرية، لأن منطقة الرفض للفرضية الصفرية في الجهة اليسرى. لذلك لا يوجد ما يكفي لتدعيم فرضية الباحث بأن التباين في اطوال النساء ذوات الاعمار من (20 إلى 25) يقل عن التباين في اطوال الرجال من نفس العمر على مستوى الدلالة $\alpha = 0.01$

دعنا نفرض أن الفرضية البديلة كانت غير متجهة



$$H_0: \sigma^2 = (2.5)^2 = 6.25$$

$$H_1: \sigma^2 \neq 6.25$$

بما ان الفرضية البديلة غير متجهة تكون منطقة الرفض على الذيلين كما يلي
القيم احرجة لمنطقة القبول $\chi^2_{40, 0.005} = 66.77$ والقيمة $\chi^2_{40, 0.995} = 20.71$
القيمة المحسوبة

$$\chi^2_{n-1} = \frac{(n-1)S^2}{\sigma^2} = \frac{(41-1)(2.133)^2}{6.25} = 29.118$$

بما ان القيمة المحسوبة 29.118 تقع ضمن منطقة قبول الفرضية الصفرية لا نستطيع رفض الفرضية الصفرية.

لذلك لا يوجد ما يكفي لتدعيم فرضية الباحث بأن التباين في اطوال النساء ذوات الاعمار 20 على 25 يختلف عن التباين في اطوال الرجال من نفس العمر على مستوى الدلالة $\alpha = 0.01$

فترة الثقة للتباين: فترة الثقة 95% تكتب

$$0.95 = \frac{(n-1)S^2}{\chi^2_{(n-1,0.975)}} \geq \sigma^2 \geq \frac{(n-1)S^2}{\chi^2_{(n-1,0.025)}}$$

تعني أن الاحتمال = 0.95 بأن تقع القيمة الصحيحة أو الحقيقية للتباين بين الحدين

$$\frac{(n-1)S^2}{\chi^2_{(n-1,0.025)}} \text{ كحد أعلى ، } \frac{(n-1)S^2}{\chi^2_{(n-1,0.975)}} \text{ كحد أدنى.}$$

مثال 3-6: افرض ان عينة من 25 شخصاً تم اختيارهم بشكل عشوائي من مجتمع يتخذ شكل التوزيع الطبيعي في الخاصية تحت الدراسة. وعند حساب التباين غير المتحيز لتلك الخاصية في أفراد العينة، وجد أنه = 16، كون فترة الثقة 95% للتباين في هذه الخاصية في المجتمع الإحصائي.

الحل:

بالرجوع إلى جدول (χ^2) نجد أن $\chi^2_{(24,0.975)} = 12.40$ وأن $\chi^2_{(24,0.025)} = 39.36$ وبذلك يكون الحد الأدنى = $\frac{16 \cdot 24}{39.36} = 9.76$

وبذلك يكون الحد الأعلى = $\frac{16 \cdot 24}{12.40} = 30.97$

توزيع ف

يستخدم لاختبار تساوي التباين في مجتمعين احصائيين

$$F = \frac{\chi^2_{\gamma_1/\gamma_2}}{\chi^2_{\gamma_2/\gamma_1}}$$

ف متغير عشوائي يتكون من النسبة لقيمتين مستقلتين من قيم χ^2 كل منهما مقسومة على درجة حريتها.

2-6 اختبار فرضية تتعلق بتساوي التباين لمجتمعين مستقلين.

Testing Hypothesis about Two Independent Variances

$$H_0: \sigma_1^2 = \sigma_2^2$$

$$F = \frac{S_1^2}{S_2^2}$$

$$S^2_2$$

$$\gamma_1 = n_1 - 1 \quad \text{درجات الحرية للبسط}$$

$$\gamma_2 = n_2 - 1 \quad \text{درجات الحرية للمقام}$$

يمكن استخدام اختبار F بذيل أو بذيلين.

الفرضية متجهة: إذا استخدم بذيل واحد تكون القيم الحرجة للرفض أو القبول مباشرة

في جدول F .

الفرضية غير متجهة: يجب أن نقسم $\alpha/2$ ونستخرج قيمتين حرجتين من قيم F هما:

$$F(\gamma_1, \gamma_2, \frac{\alpha}{2}) \quad \text{كحد أدنى}$$

$$F(\gamma_1, \gamma_2, 1 - \frac{\alpha}{2}) \quad \text{الحد الأعلى ويستخرج مباشرة من الجدول}$$

$$\frac{1}{F(\gamma_2, \gamma_1, 1 - \frac{\alpha}{2})} = F(\gamma_1, \gamma_2, \frac{\alpha}{2}) \quad \text{الحد الأدنى يستخرج من العلاقة}$$

مثال 4-6) اعتقد باحث أن الطلاب الاذكياء يبدعون في التحصيل في جو تنافسي بينما يقل تحصيلهم نسبياً في جو تعاوني وبذلك يزداد التباين في التحصيل اذا تم التدريس في جو تنافسي ولكنه يقل إذا تم تدريسهم في جو تعاوني. ولأختبار هذه الفرضية قام باختيار مجموعتين بشكل عشوائي درست الأولى في جو تنافسي والثانية في جو تعاوني. وعند انتهاء التجربة استخرج الانحراف غير المتحيز لهاتين المجموعتين فكان لمجموعة الجو التنافسي $S_1=6.4$ وللمجموعة الجو التعاوني $S_2=3.9$ فإذا كان عدد أفراد المجموعة الاولى $n_1=17$ وعدد أفراد المجموعة الثانية $n_2=15$ فهل تدعم هذه البيانات صحة إدعاء الباحث؟

وإذا كان ادعاؤه هو ان اختلاف الجو التدريسي لا يؤثر على التباين في التحصيل، فهل

تدعم هذه البيانات صحة مثل هذا الادعاء؟ استخدم في الحالتين $\alpha=0.05$

في حالة الفرضية متجهة:

$$H_0: \sigma_1^2 = \sigma_2^2$$

$$H_1: \sigma_1^2 > \sigma_2^2$$

بما ان البيانات مستقلة إذا الاختبار المناسب هو F

$$\gamma_1 = n_1 - 1 = 17 - 1 = 16 \quad \text{درجات الحرية للبسط}$$

$$\gamma_2 = n_2 - 1 = 15 - 1 = 14 \quad \text{درجات الحرية للمقام}$$

$$F = \frac{S_1^2}{S_2^2} = \frac{(6.4)^2}{(3.9)^2} = 2.69 \quad \text{نستخرج قيمة المختبر الإحصائي}$$

القيمة الحرجة F هي $F(16,14,0.95) = 2.44$

بما أن القيمة المحسوبة ($F=2.69$) أكبر من القيمة الحرجة ($F(16,14,0.95) = 2.44$)
لذلك نرفض الفرضية الصفرية ونستنتج أن تدريس الطلاب في جو تنافسي يؤدي الى جعل
التسابقين في علاماتهم أعلى وبدلالة احصائية غير مستوى $\alpha=0.05$ مما هي عليه الحال في الجو
التعاوني.

في حالة الفرضية غير متجهة:

$$H_0: \sigma_1^2 = \sigma_2^2$$

$$H_1: \sigma_1^2 \neq \sigma_2^2$$

القيمة المحسوبة ($F=2.69$)

$$F = \frac{S_1^2}{S_2^2} = \frac{(6.4)^2}{(3.9)^2} = 2.69 \quad \text{نستخرج قيمة المختبر الإحصائي}$$

$$F = \frac{S_2^2}{S_1^2} = \frac{(3.9)^2}{(6.4)^2} = 0.37 \quad \text{نستخرج قيمة المختبر الإحصائي}$$

القيمتين الحرجتين للمختبر الاحصائي F وهما:

$$F(\gamma_1, \gamma_2, 1 - \frac{\alpha}{2}) = F(16,14,0.975) = 2.95 \quad \text{الحد الأعلى يستخرج مباشرة من الجدول}$$

الحد الأدنى يستخرج من العلاقة

$$\frac{1}{F(\gamma_2, \gamma_1, 1 - \frac{\alpha}{2})} = F(\gamma_1, \gamma_2, \frac{\alpha}{2}) = F(16,14,0.025) = 1 / F(14,16,0.975) = 1/2.75 = 0.36$$

بما أن القيمة المحسوبة لـ F هي: 0.37 و 2.6 تقع ضمن منطقة القبول بين الحدين 2.95 و 0.36 ولذلك نقبل الفرضية الصفرية ونستنتج أن تدريس الطلاب في جو تنافسي لا يؤدي إلى جعل الاختلاف في التباين في علاماتهم عن التباين في علامات الذين يدرسون في جو تعاوني ذا دلالة احصائية.

3-6 اختبار فرضية تتعلق بتساوي التباين لمجتمعين غير مستقلين.

Testing Hypothesis about Two Dependent Variances

$$H_0: \sigma_1^2 = \sigma_2^2$$

$$\gamma = n - 2 \quad \text{درجات الحرية}$$

$$T = \frac{S_1^2 - S_2^2}{\sqrt{\frac{4S_1^2 - S_2^2}{n-2} * (1-r^2)}}$$

S_1^2 : التباين غير المتحيز للعينة الأولى.

S_2^2 : التباين غير المتحيز للعينة الثانية.

n : عدد أزواج المشاهدات.

r : معامل الارتباط بين المشاهدات في العينة الأولى والمشاهدات في العينة الثانية.

مثال 5-6: اعتقد باحث أن التحانس في الأداء على مقياس للالتجاهات نحو المدرسة

يزداد مع تقدم الطالب في المستوى التعليمي. وبمعنى آخر يقل التباين كلما ارتفع الطالب من مستوى تعليمي لآخر. وللتحقق من صحة هذا الادعاء أجرى اختباراً للالتجاهات على عينة من الطلاب من مستوى الثالث الاعدادي ثم أعاد إجراء الاختبار على نفس العينة وهم في مستوى الثالث الثانوي. وقد كان عدد الذين اخذوا الاختبار في المراتين يساوي 74 طالباً. وكان التباين في العلامات في الإجراء الأول $S_1^2 = 105$ وفي الإجراء الثاني أصبح التباين $S_2^2 = 82$ ، أما معامل الارتباط بين علامات الطلاب في الإجراءين فكان $r = 0.8$ هل تدعم هذه البيانات صحة إدعاء الباحث؟ استخدم $\alpha = 0.05$

الحل:

$$H_0: \sigma_1^2 = \sigma_2^2$$

$$H_1: \sigma_1^2 > \sigma_2^2$$

 σ_1^2 : التباين في الثالث الإعدادي.

 σ_2^2 : التباين في الثالث الثانوي.

القيمة المحسوبة للمختبر الاحصائي T

$$T = \frac{S_1^2 - S_2^2}{\sqrt{\frac{4 S_1^2 S_2^2 * (1 - r^2)}{n - 2}}}$$

$$T = \frac{105.82}{\sqrt{\frac{4 * 105 * 82 * (1 - 0.8^2)}{74 - 2}}} = 1.752$$

القيمة الحرجة $T_{\gamma, 1-\alpha}$

$$\gamma = 74 - 2 = 72 \text{ درجات الحرية}$$

الاختبار بذييل واحد One Tail

$$\alpha = 0.05 \text{ مستوى الدلالة}$$

$$T_{\gamma, 1-\alpha} = T_{72, 1-0.05} = T_{72, 0.95} = 1.671$$

النتيجة:

بما أن القيمة المحسوبة للمختبر الاحصائي T (1.752) أكبر من القيمة الحرجة (1.671) نرفض الفرضية الصفرية على مستوى $\alpha = 0.05$ ، ونستنتج أن البيانات تدل على أن التباين في إتحادات الطلاب نحو المدرسة يقل مع ارتفاع مستواهم التعليمي.

4-6 تمارين Exercise

تمرين (1) : ضع رمز الإجابة (نعم أو لا) في المربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										

1- يمكن إهمال افتراض تجانس التباين في اختبار T إذا تساوى عدد الأفراد في كل من المجموعتين.

2- يتم اختبار الفرضية $H_0: \sigma^2 = C$ باستخدام اختبار T

3- يتم اختبار الفرضية $H_0: \sigma_1^2 = \sigma_2^2$ مشاهدات مستقلة باستخدام اختبار F

4- يتم اختبار الفرضية $H_0: \sigma_1^2 = \sigma_2^2$ مشاهدات غير مستقلة باستخدام اختبار T

5- تصاغ الفرضيات الاحصائية بعد أن يتم جمع البيانات.

6- اختبار التباين بصورة عامة لا تتأثر بعدم الإيفاء بافتراضات التوزيع الطبيعي.

7- إذا كانت الفرضية البديلة $H_1: \sigma_1^2 > \sigma_2^2$ فإن القيمة الحرجة للاحصائي F هي التي

يعطيها جدول F مباشرة؟

8- إذا كانت $n_1 = n_2$ فلماذا نختتم باختبار الفرضية $H_0: \sigma_1^2 = \sigma_2^2$ ؟

تمرين (2) : ضع رمز الإجابة الصحيحة في المربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										
الرقم	11	12	13	14	15	16	17	18	19	20
الإجابة										

1- أن توزيع χ^2 يتصف بما يلي:

أ- توقع التوزيع = درجات الحرية.

ب- يكون دائماً متماثلاً.

ج- يكون ملتوياً التواء موجباً.

د- يكون ملتوياً التواء سالباً.

2- أن توزيع T يتصف بما يلي:

- أ- توقع التوزيع = درجات الحرية.
 ج- يكون ملتوياً التواء موجباً.
 ب- يكون دائماً متماثلاً.
 د- يكون ملتوياً التواء سالباً.

3- أن توزيع F يتصف بما يلي:

- أ- توقع التوزيع = درجات الحرية.
 ج- يكون ملتوياً التواء موجباً.
 ب- يكون دائماً متماثلاً.
 د- يكون ملتوياً التواء سالباً.

4- إن القيمة الحرجة ($F_{5,7,0.05}$) هي:

- أ- 3.95 ب- 7.46 ج- 4.88 د- 16.45

5- إن القيمة الحرجة ($\chi^2_{5,0.99}$) هي:

- أ- 0.554 ب- 0.752 ج- 1.145 د- 1.610

6- إن القيمة الحرجة ($T_{24,0.975}$) هي:

- أ- 2.060 ب- 2.069 ج- 2.064 د- 1.711

7- يتم اختبار الفرضية $H_0: \sigma^2 = C$ باستخدام اختبار:

- أ- Z ب- T ج- F د- χ^2

8- يتم اختبار الفرضية $H_0: \sigma_1^2 = \sigma_2^2$ مشاهدات مستقلة باستخدام اختبار.....

- أ- Z ب- T ج- F د- χ^2

9- عينتان مستقلتان، التباين في المشاهدات في أحد المتغيرات في الأولى = 100 وعدد أفرادها =

15 والتباين في المشاهدات في نفس المتغير للعينة الثانية = 64 وحجمها = 36، هل يختلف

التباين في المجتمعين احصائياً؟

أ- لا، لأن $F=1.56$ وهي ذات دلالة احصائية.

ب- نعم، لأن $F=1.56$ وهي ذات دلالة احصائية.

ج- نعم، لأن $F=1.56$ وهي ليست ذات دلالة احصائية.

د- لا، لأن $F=1.56$ وهي ليست ذات دلالة احصائية.

10- يمكن إهمال افتراض تجانس التباين في اختبار T إذا :

- أ- عدد الأفراد في المجموعة الأولى يساوي عدد الأفراد في المجموعة الثانية.
- ب- عدد الأفراد في المجموعة الأولى أكبر من عدد الأفراد في المجموعة الثانية
- ج- عدد الأفراد في المجموعة الأولى أقل من عدد الأفراد في المجموعة الثانية.
- د- لا شيء مما ذكر.

11- أي من التالية ليس من خصائص توزيع كاي تربيع (χ^2)

- أ- يأخذ قيماً سالبة.
- ب- القيمة العظمى للاحتمال فيه (68% من الحالات) تنحصر بين 0 - 1.
- ج- ملتو التواء موجب، وإن الالتواء يقل بازدياد درجات الحرية.
- د- ليس له شكل محدد ولكن يعتمد شكله على درجات الحرية.

12- يتم اختبار الفرضية $H_0: \sigma_1^2 = \sigma_2^2$ مشاهدات غير مستقلة باستخدام اختبار.....

- أ- Z
- ب- T
- ج- F
- د- χ^2

13- في توزيع χ^2 إن القيمة العظمى للاحتمال التي تنحصر بين 0 - 1 هي:

- أ- 0.34
- ب- 0.68
- ج- 0.95
- د- 0.99

14- إذا كان الانحراف المعياري للمجتمع = 2.5، وكان الانحراف المعياري لعينة = 2.13

وعدد العينة = 31 ومستوى الدلالة = 0.01 فإن قيمة χ^2 المحسوبة تساوي:

- أ- 0.34
- ب- 0.68
- ج- 0.95
- د- 0.99

15- إذا كان الانحراف المعياري للمجتمع = 2.5، وكان الانحراف المعياري للعينة = 2.13

وعدد العينة = 31 ومستوى الدلالة = 0.01 فإن قيمة χ^2 الحرجة تساوي:

- أ- 20.599
- ب- 16.306
- ج- 18.493
- د- 14.953

الفصل السابع

اختبار الفرضيات

حول معاملات الارتباط

Hypotheses Testing Regarding Correlation Coefficients

- 1-7 مقدمة
- 2-7 اختبار الفرضية حول معامل ارتباط واحد.
- 3-7 اختبار الفرضيات حول الفرق بين معاملي ارتباط مستقلين.
- 4-7 اختبار الفرضيات حول معاملي ارتباط للبيانات غير المستقلة.
- 5-7 استخدام برمجية SPSS في حل المسائل.
- 6-7 تمارين Exercise .

الفصل السابع

اختبار الفرضيات حول معاملات الارتباط

Hypotheses Testing Regarding Correlation Coefficients

1-7 مقدمة

الارتباط Correlation

إن نظرية الارتباط تظهر قوة العلاقة بين متغيرين مع إمكانية تحديد نوع وقوة العلاقة بين الظواهر، كالعلاقة بين مستوى التعليم والأداء، والعلاقة بين معدل الثانوية العامة ومعدل الجامعة، والعلاقة بين المستوى الاقتصادي والتحصيل.

إن الهدف من تحليل الارتباط Correlation هو معرفة وجود علاقة بين متغيرين أو مجموعة من المتغيرات المستقلة Independent Variables (X_i) مع المتغير التابع Variable Dependent (Y_i) من عدم وجودها، وهناك عدة مقاييس لتحديد درجة العلاقة والارتباط بين المتغيرات.

– قياس الارتباط Measures of Correlation

1- رسم شكل الانتشار Scatter Plot.

2- القياس الكمي للارتباط Quantitative Measure

معامل الارتباط الخطي البسيط Simple Linear Correlation

أ- معامل ارتباط بيرسون الخطي Person Linear Correlation Coefficient

معامل ارتباط بيرسون يقيس قوة واتجاه العلاقة الخطية فقط بين متغيرين كميين.

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

$$r_{xy} = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}$$

$$r_{xy} = \frac{\sum Z_x Z_y}{n} \quad \frac{\sum \text{نفس نفس}}{n} = r$$

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n(\sum x^2) - (\sum x)^2] [n(\sum y^2) - (\sum y)^2]}}$$

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n(\sum x^2) - (\sum x)^2] [n(\sum y^2) - (\sum y)^2]}}$$

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{n \sigma_x \sigma_y}$$

$$r = \frac{\sum xy - n \bar{x} \bar{y}}{\sqrt{[\sum x^2 - n \bar{x}^2] [\sum y^2 - n \bar{y}^2]}}$$

$$r = \frac{(\sum xy) - n \bar{X} \bar{Y}}{[\sqrt{(\sum x^2) - n \bar{X}^2}] [\sqrt{(\sum y^2) - n \bar{Y}^2}]}$$

* الشروط الواجب توافرها لاستخدام معامل ارتباط بيرسون

1. وجود علاقة خطية بين المتغيرين.
2. يجب أن تكون العينة عشوائية وقيم أفراد العينة مستقلة عن بعضها البعض.

*** تقييم قيمة معامل الارتباط Correlation Coefficient Evaluation**

جدول (24): تقييم قيمة معامل الارتباط حسب تصنيف

(Hinkle and Others, 1979)

التفسير	الفترة
منخفض جداً	من 0.00 - أقل من 0.30
منخفض	من 0.30 - أقل من 0.50
متوسط	من 0.50 - أقل من 0.70
عال	من 0.70 - أقل من 0.90
عال جداً	من 0.90 - أقل من 1.00

* (تحليل الخليلي) تصنيف (Hinkle and Others, 1979)

ب- معامل ارتباط سبيرمان للترتيب Spearman Correlation Coefficient

يستخدم معامل ارتباط سبيرمان ومعامل كندال تاو لقياس قوة الارتباط بين متغيرين

ترتيبين Ordinal.

$$r = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

$$r = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

حيث: d تعني الفرق بين رتبة X ورتبة Y

ج- معامل ارتباط إيتا Eta η بين متغيرين كل منهما متصل والعلاقة بينهما انحنائية.

د- معامل التوافق (م ت) Contingency Coefficient (CC)

معامل ارتباط بين متغيرين كل منهما منفصل، ولكن ليس بالضرورة أن يكون أي منهما منفصلاً ثنائياً، يمكن أن يستخدم عندما تكون عدد الفئات في أحد المتغيرين أو كليهما اثنين أو أكثر، معامل ارتباط بين متغيرين أحدهما أو كلاهما ينقسم إلى أكثر من حالتين، وقد يكون أحد أو كلا المتغيرين وصفية.

هـ- راسيرال رتي (رد) Rank Biserial (r_{rb})

معامل ارتباط بين متغيرين أحد هذين المتغيرين ثنائي منفصل والآخر متغير رتي. مثل علاقة نجاح أو فشل الطالب في دراسته والمستوى الاقتصادي لولي الأمر.

و- تتراشورك (دت) Tetrachoric (r_t)

معامل ارتباط بين متغيرين كل منهما منفصل بالتحويل.

ز- معامل الاقتران (r_A) Coefficient of Association :

يستخدم عندما تكون بيانات كلا المتغيرين X, Y أو أحدهما غير قابلة للترتيب التصاعدي أو التنازلي، وكان عدد الحالات التي فيها كل من المتغيرين هي حالتين فقط.

ح- معامل ارتباط راسيرال (رب) Biserial (r_b)

معامل ارتباط بين متغيرين أحد هذين المتغيرين يقع على مقياس فتوي أو مقياس نسبي والآخر متغير ثنائي منفصل ولكن بصورة غير طبيعية.

ط- معامل ارتباط بوينت راسيرال (ردد) Point Biserial (r_{pb})

معامل ارتباط بين متغيرين بحيث يكون أحد المتغيرين منفصلاً ثنائياً بصورة طبيعية مثل مستغير الجنس والمتغير الثاني يقع على مقياس فتوي Interval أو مقياس نسبي Ratio مثل متغير الذكاء.

$$r_{pb} = [(Y'_2 - Y'_1) / S_y] (\sqrt{pq})$$

ي- معامل ارتباط فاي (Φ) Phi

معامل ارتباط بين متغيرين كل منهما منفصل ثنائي بصورة طبيعية Dichotomous، ويقع هذين المتغيرين على مقياس اسمي (Nominal).
يحسب معامل ارتباط فاي (Φ) بالمعادلة التالية:

$$\Phi = \frac{P_{1xy} - P_{1x} P_{1y}}{\sqrt{P_{1x} P_{0x} P_{1y} P_{0y}}}$$

P_{1xy} : نسبة الأفراد أصحاب العلامة 1 على المتغيرين.

P_{1x} : نسبة الأفراد أصحاب العلامة 1 على المتغير x .

$P1_y$: نسبة الأفراد أصحاب العلامة 1 على المتغير y .

$P0_x$: نسبة الأفراد أصحاب العلامة 0 على المتغير x .

$P0_y$: نسبة الأفراد أصحاب العلامة 0 على المتغير y .

معامل التحديد **Determinant Coefficient** التباين المفسر أو المشترك r^2

* أراد احسد الباحثين أن يدرس العلاقة بين الذكاء C ، والابتكار IQ فحسب معامل

ارتباط بيرسون بين أداء الطلاب على اختبار للذكاء C واختبار للابتكار IQ فكان يساوي $r_{q,c}$

$$r^2 = r * r \quad , \quad = 0.7$$

2-7 اختبار الفرضية حول معامل ارتباط واحد.

لمعرفة ما اذا كانت معاملات الارتباط التي نستخرجها بين متغيراتنا ذات دلالة احصائية

ام لا .

ما المقصود بعبارة "ان معامل الارتباط بين متغيرين كان ذا دلالة احصائية"؟

ان هذه العبارة استدلال احصائي ينتج عن رفض الفرضية الصفرية $H_0: \rho = 0$ وهي

تعني ان معامل الارتباط في المجتمع بين المتغيرين لا يختلف عن صفر، ورفض هذه الفرضية

الصفرية يعني ان قيمة معامل الارتباط في المجتمع تختلف عن الصفر اختلافاً جوهرياً، أي أنها لم

تنشأ عن خطأ المعاينة.

يستخدم توزيع المعاينة لمعامل الارتباط بيرسون r في اختبار الفرضية الصفرية $H_0: \rho = 0$

يتخذ شكل توزيع T بدرجات حرية $\gamma = n - 2$ ، حيث n هي حجم العينة.

$$T = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$

مثال 14-1: وجد في احدى الدراسات ان معامل الارتباط بين المستوى الاقتصادي

الاجتماعي والتحصيل $= 0.4$ عند استخدام عينة عشوائية حجمها - 30 طالب، فهل هذا

المعامل يختلف جوهرياً عن الصفر؟

$$H_0: \rho = 0$$

$$H_1: \rho \neq 0$$

$$T = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{0.4}{\sqrt{\frac{1-0.4^2}{30-2}}} = \frac{0.4}{\sqrt{\frac{0.84}{28}}} = \frac{0.4}{0.1732} = 2.31$$

$$T_{28,0.975} = 2.048 \text{ القيمة الحرجة}$$

النتيجة: نرفض الفرضية الصفرية لأن القيمة الحرجة للمختبر الإحصائي 2.048 أقل من القيمة المحسوبة 2.31، ونقول أن معامل الارتباط بين المستوى الاقتصادي الاجتماعي والنحصيل 0.4 كان ذا دلالة إحصائية، أي يختلف جوهرياً عن الصفر؟ يمكن تحديد القيمة الحرجة لمعامل الارتباط الذي نرفض أو نقبل الفرضية الصفرية وهي:

$$r_c = \frac{T}{\sqrt{T^2 + \gamma}}$$

$$\gamma = n - 2$$

القيمة الحرجة للرفض أو القبول حسب درجات الحرية ومستوى الدلالة T

استخدام جداول r_c مباشرة دون الحاجة لحساب T

عندما $r=0.4$ يسدل على ارتباط ذي دلالة احصائية على مستوى $\alpha=0.05$ والقيمة

الحرجة في الجدول هي 0.374

3-7 اختبار الفرضيات حول الفرق بين معاملي ارتباط مستقلين. Hypotheses Testing Regarding the Difference between two Independent Correlation Coefficients

$$H_0: \rho_1 = \rho_2$$

نستخدم اختبار Z بعد الاستعانة بعلامات فشر المعيارية.

$$Z = \frac{F_1 - F_2}{\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}}$$

حيث:

n_1 : حجم العينة الأولى.

F_1 : علامة فشر المقابلة لمعامل الارتباط في العينة الأولى.

n_2 : حجم العينة الثانية.

F_2 : علامة فشر المقابلة لمعامل الارتباط في العينة الثانية.

مثال 14-2: اعتقد باحث أن قوة العلاقة بين القدرة القرائية والتحصيل في مجتمع الإناث في جامعة الاسراء أقل منها في مجتمع الذكور في نفس المستوى. وعندما استخرج معامل الارتباط بين هذين المتغيرين على عينة عشوائية من الإناث من هذه الفئة مؤلفة من 40 طالبة وجده مساوياً 0.7، وبالمقابل وجد معامل الارتباط بين المتغيرين مساوياً 0.8 عندما استخرجه من خلال عينة من الذكور مؤلفة من 50 طالباً. فهل تدعم هذه البيانات صحة ادعاء الباحث؟

$$H_0: \rho_1 = \rho_2$$

$$H_1: \rho_1 > \rho_2$$

حيث:

ρ_1 : معامل الارتباط بين القدرة القرائية والتحصيل في مجتمع الذكور.

ρ_2 : معامل الارتباط بين القدرة القرائية والتحصيل في مجتمع الإناث.

وحيث أن $r_1 = 0.8$ فإن $F_1 = 1.099$

وكذلك $r_2 = 0.7$ فإن $F_2 = 0.867$

$$Z = \frac{F_1 - F_2}{\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}} = \frac{0.867 - 1.099}{\sqrt{\frac{1}{40 - 3} + \frac{1}{50 - 3}}} = -1.06$$

النتيجة:

القيمة الحسرة للمختبر الاحصائي Z على مستوى $\alpha = 0.05$ تساوي -1.64 ومن الواضح أن القيمة المحسوبة للمختبر الاحصائي واقعة في منطقة قبول الفرضية الصفرية على هذا المستوى وبذلك نقبل الفرضية الصفرية، ونستدل على أن البيانات المتوفرة لا تدل على أن قوة العلاقة بين القدرة القرائية والتحصيل هي أقل في مجتمع الإناث عما هي عليه في مجتمع الذكور.

4-7 اختبار الفرضيات حول معاملي ارتباط للبيانات غير المستقلة.

Hypotheses Testing Regarding two dependent Correlation Coefficients

مثال: صدق التنبؤ بمعدل الطالب في نهاية المرحلة الجامعية من خلال معدله في اثنائية العامة مقابل صدق التنبؤ من خلال معدله في المدرسة على نفس افراد العينة.

$$H_0: \rho_{31} = \rho_{32}$$

$$H_1: \rho_{31} > \rho_{32}$$

ρ_{31} : معامل الارتباط في المجتمع بين المتغير 3 والمتغير 1

ρ_{32} : معامل الارتباط في المجتمع بين المتغير 3 والمتغير 2

$$T = (r_{31} - r_{32}) \sqrt{\frac{(n-3)(1+r_{12})}{2(1-r_{31}^2 - r_{32}^2 - r_{12}^2 + 2r_{31}r_{32}r_{21})}}$$

$$\gamma = n - 3$$

حيث:

n : حجم العينة.

r_{31} : معامل الارتباط بين المتغيرين 3، 1 في العينة.

r_{32} : معامل الارتباط بين المتغيرين 3، 2 في العينة.

r_{12} : معامل الارتباط بين المتغيرين 1، 2 في العينة.

مثال 14-3: اعتقد باحث أن التحصيل يرتبط بدرجة أقوى بالقدرة القرائية منه بالقدرة الكتابية في مستوى الابتدائي. ولأجل ذلك طور اختبارين أحدهما يقيس القدرة القرائية والآخر يقيس القدرة الكتابية ثم طبقهما على عينة مؤلفة من 67 طالباً. وفي نهاية الفصل الدراسي سجل معدل هؤلاء في المدرسة كمؤشر للتحصيل. ولدى حساب معاملات الارتباط بين هذه المتغيرات حصل على القيم التالية:

r_{31} معامل الارتباط بين التحصيل والقدرة القرائية = 0.8

r_{32} معامل الارتباط بين التحصيل والقدرة الكتابية = 0.65

r_{12} معامل الارتباط بين القدرة القرائية والقدرة الكتابية = 0.5

فهل تدل البيانات على صحة ادعاء الباحث؟

المتغير المشترك 3 هو التحصيل، المتغير 1 القدرة القرائية، المتغير 2 القدرة الكتابية.

$$H_0: \rho_{31} = \rho_{32}$$

$$H_1: \rho_{31} > \rho_{32}$$

$$T = (r_{31} - r_{32}) \sqrt{\frac{(n-3)(1+r_{12})}{2(1-r_{31}^2 - r_{32}^2 - r_{12}^2 + 2r_{31}r_{32}r_{21})}}$$

$$T = (0.8 - 0.65) \sqrt{\frac{(67-3)(1+0.5)}{2(1-(0.8)^2 - (0.65)^2 - (0.5)^2 + 2*0.8*0.65*0.5)}} = 2.28$$

$$\gamma = n - 3 = 67 - 3 = 64$$

$$T_{64,0.95} = 1.67$$

نرفض الفرضية الصفرية على مستوى $P < 0.05$ ونستنتج أن البيانات لا تدل على عكس ما يدعيه الباحث. إذا كان معامل الارتباط بين التحصيل والقدرة القرائية أعلى منه بين التحصيل والقدرة الكتابية.

الأساليب الإحصائية المناسبة لدراسة العلاقة وفقاً لعدد المتغيرات.

أساليب دراسة العلاقة بين متغيرين	مجموعة من المتغيرات المستقلة ومتغير تابع واحد	العلاقة بين مجموعة من المتغيرات المستقلة ومجموعة من المتغيرات التابعة	العلاقة بين متغيرين مع ضبط الثالث
Person's Correlation Rank- Correlation Sperman Rho Kendall's tau Biserial Correlation Widespread Biserial Correlation Point-Biserial Correlation Tetrachoric Correlation Phi Coefficient Contingency Coefficient Correlation ratio	Multiple Linear regression Discriminate Function	Canonical Correlation	Partial Correlation Part correlation

أساليب حساب العلاقة المناسبة وفقاً لمستوى القياس للمتغيرين:

المتغير الأول	المتغير الثاني	المقاييس المناسبة
فئري أو سبي	فئري أو نسبي	معامل بيرسون (Pearson product Moment Correlation, r.) Coefficient — حاصل ضرب العزوم — (إذا كانت العلاقة خطية). نسبة الارتباط Correlation Ratio (معامل أيتا) (إذا كانت العلاقة غير خطية)

<p>معامل سبيرمان لارتباط الرتب Spearman's Rank</p> <p>معامل كادال تسو لارتباط الرتب Kedall's Tau Rank</p> <p>معامل الارتباط لقياس الارتباط ووزن الرتب بميزان فترتي.</p> <p>معامل الارتباط لقياس الارتباط مع عدم وزن الرتب بميزان فترتي.</p> <p>معامل الارتباط لجودمان وكروسكال Goodman and</p> <p>معامل الارتباط لقياس الارتباط Kruskal's Coefficient of Ordinal Association</p>	رتبي	رتبي
<p>معامل التباين المتماثل لختمان</p> <p>معامل فاي ، معامل الارتباط ليول، معامل التجمع ليول (عندما يشمل كل متغير على قسمين).</p> <p>معامل الارتباط لبرسون.</p> <p>معامل التوافق -التصاحب- Contingency عندما يكون أحد المتغيرين أو كليهما متعدد الفئات.</p> <p>معامل تراشورك Tetrachoric Coefficient : يستخدم إذا كان مستوى القياس في المتغيرين متصلًا ثم حولت إلى اسمية.</p>	اسمي	اسمي
<p>معامل وليكوكسون للارتباط (إذا لم يكن هناك تمييز بين المتغير المستقل والتابع).</p> <p>معامل وليكوكسون لإشارات الرتب إذا كان هناك تمييز بين المتغير المستقل والتابع. (هناك معامل خاص عندما يتكون المتغير الاسمي من قسمين، و آخر عندما يتكون المتغير الاسمي من أكثر من قسمين).</p> <p>رتب بايسريال Rank Biserial</p>	رتبي	اسمي
<p>نسبة الارتباط (مع افتراض التوزيع الاعتيادي للبيانات، وان يكون المتغير التابع هو المتغير الفترتي).</p> <p>بوينت بايسريال Point Biserial Correlation</p> <p>بايسريال Biserial Correlation (عندما يكون المتغير الاسمي أصلاً متصلًا ولكنه حول إلى اسمي كتحويل درجة مفهوم ذات سالب و موجب والتعامل معها كمتغير اسمي أو ثنائي..)</p>	فترتي أو نسبي	اسمي
<p>معامل الارتباط المتسلسل المتعدد لجاسين Jaspem Coefficient of</p> <p>Multi-serial Correlation (شرط اعتبار المتغير الرتبي متغير متصل يأخذ التوزيع الاعتيادي).</p> <p>معامل الارتباط الثنائي المتسلسل Biserial Correlation</p>	فترتي أو نسبي	رتبي

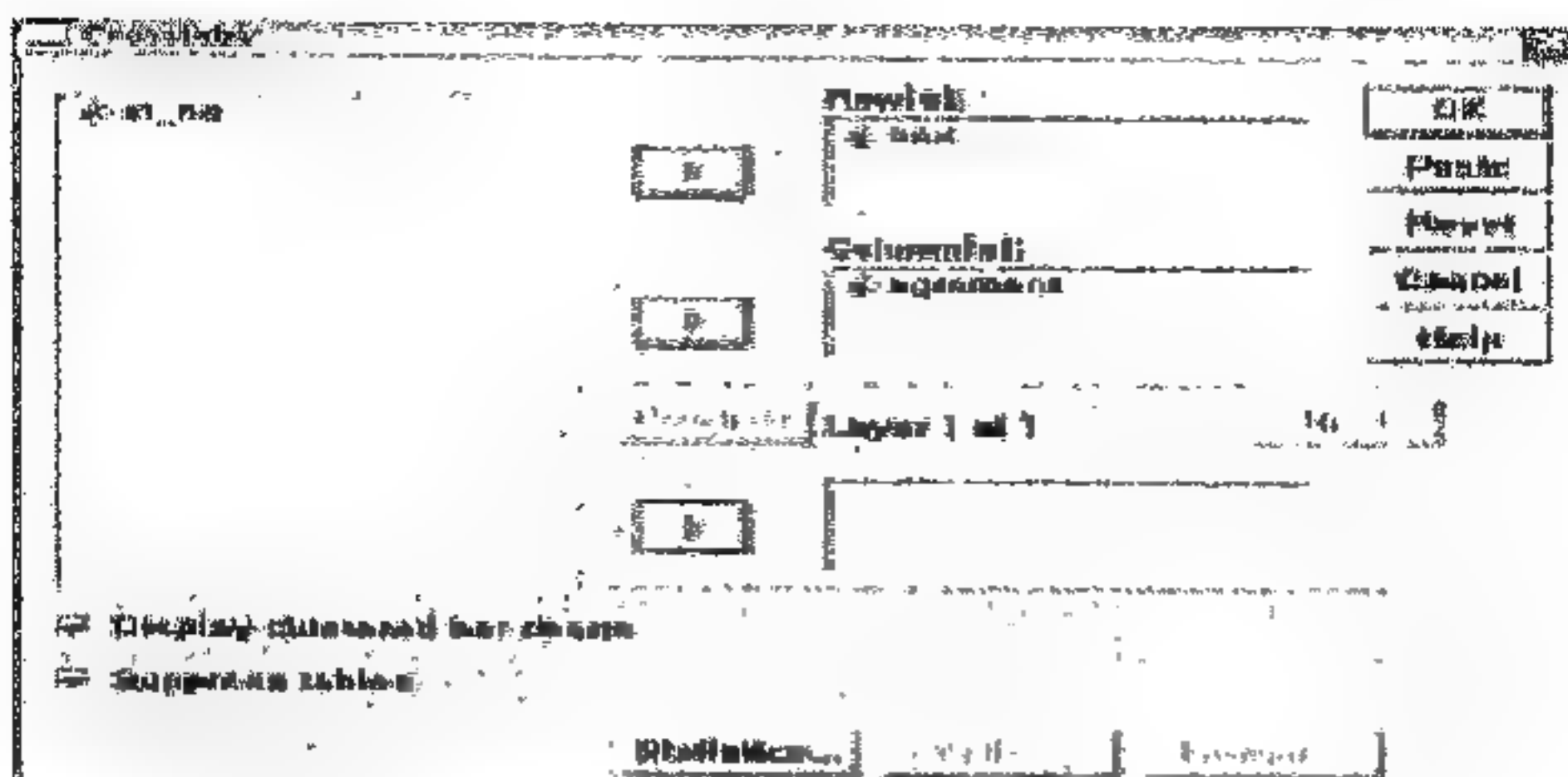
معامل فاي . معامل الارتباط الثنائي المتسلسل (عندما يكون المتغير الثنائي غير حقيقي ويكون المطلوب تقدير معامل الارتباط كما لو كان المتغير متصلًا). معامل ارتباط بيرسون (عندما يكون المتغير الثنائي متغيرًا حقيقيًا).	فئوي أو نسي	ثنائي (0-1) Dichotomous
معامل الارتباط الثنائي المتسلسل Point Biserial Coefficient Correlation (عندما يكون المتغير الثنائي غير حقيقي ويكون المطلوب تقدير معامل الارتباط كما لو كان المتغير متصلًا). معامل الارتباط الرباعي الحقيقي (فاي) Fourfold Phi Correlation (الثنائية غير حقيقية واعتبارها متصلًا). معامل الارتباط الرباعي Tetrachoric Correlation معامل ارتباط بيرسون (عندما يكون المتغير الثنائي متغيرًا حقيقيًا).	ثنائي	ثنائي
الانحدار الخطي (عند التمييز بين المتغير المستقل والتابع، العلاقة خطية، الهدف التنبؤ).	فئوي	فئوي
الانحدار المنحني (عند التمييز بين المتغير التابع والمستقل، العلاقة غير خطية، الهدف التنبؤ).		
نسبة الارتباط (عندما لا يكون هناك تمييز بين المتغير المستقل والتابع، علاقة غير خطية، ليس الاقتران هدفًا لقياس).		

5-7 استخدام برمجية SPSS في حل المسائل.

مثال 9: إذا كان غرض الباحث هو حساب معامل الارتباط بين جنس المفحوص (sex) ونسوع الإجابة عن سؤال من بدلين (Agreement) لمجموعة من 10 طلاب من الجنسين، وكانت الإجابات كما يلي:

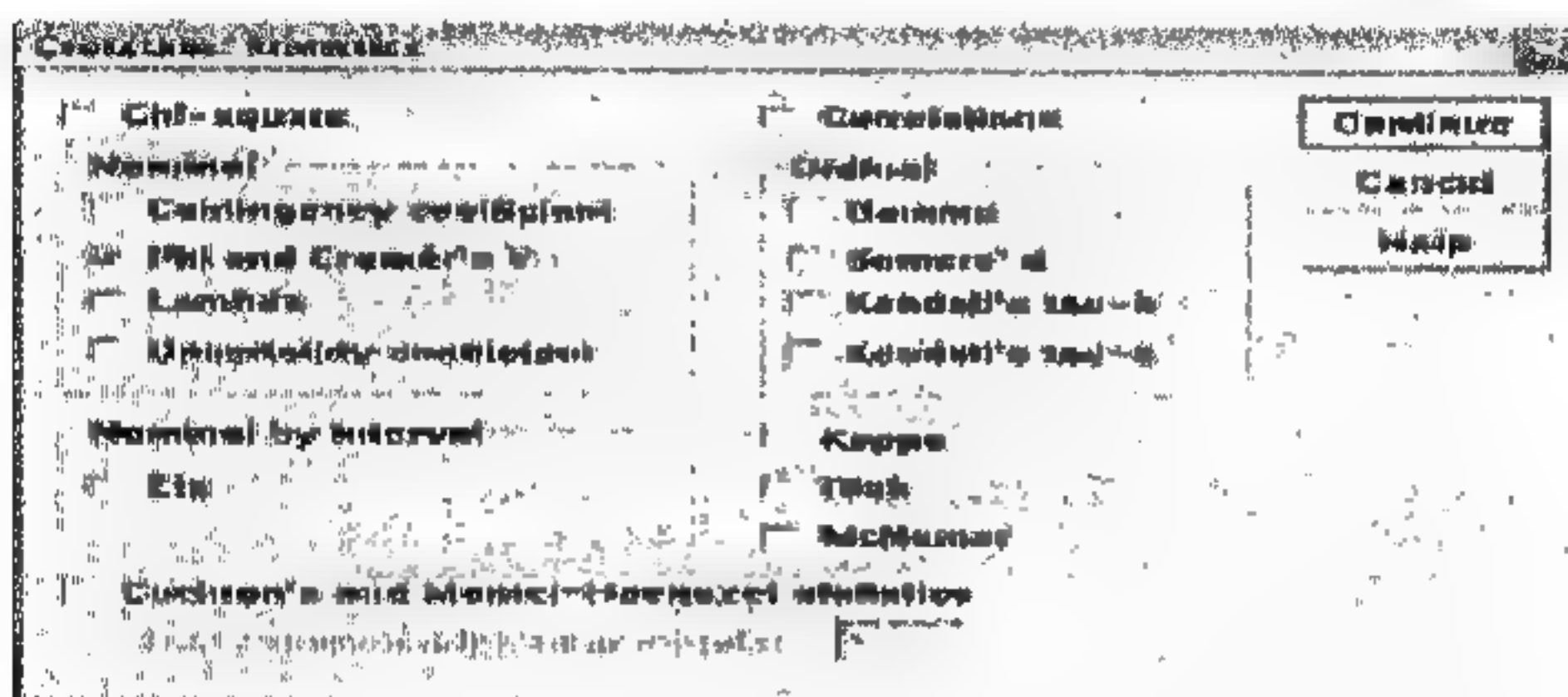
	st no	sex	agreement
1	1	1	1
2	2	1	1
3	3	1	0
4	4	0	1
5	5	0	0
6	6	1	0
7	7	1	1
8	8	1	1
9	9	0	0
10	10	0	0

Analyze – Descriptive Statistics – Crosstabs...



ضع متغير sex في نافذة Row(s) وضع متغير agreement في نافذة Column(s) ثم

اضغط زر Statistics تظهر الشاشة أدناه:



نختار Phi and Cramer's V ثم نضغط زر Continue ثم زر Ok فتظهر شاشة

المخرجات أدناه:

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
SEX * AGREEMENT	10	100.0%	0	.0%	10	100.0%

Symmetric Measures

		Value	Approx Sig
Nominal by Nominal	Phi	.408	.187
	Cramer's V	.408	.187
N of Valid Cases		10	

- a. Not assuming the null hypothesis.
- b. Using the asymptotic standard error assuming the null hypothesis.

التعليق: يشير معامل الارتباط إلى وجود علاقة ايجابية بين جنس الطالب ونوع الإجابة، حيث تفسر أن الطلاب الذكور يميلون إلى الموافقة نحو فكرة معينة بينما تميل الإناث إلى عدم الموافقة.

مثال (9-1) إذا علمت أن علامة طالب في مادة الإحصاء Stat هي X وعلامته في مادة الرياضيات Math هي y . فهل تعتقد بوجود علاقة بين العلامتين؟

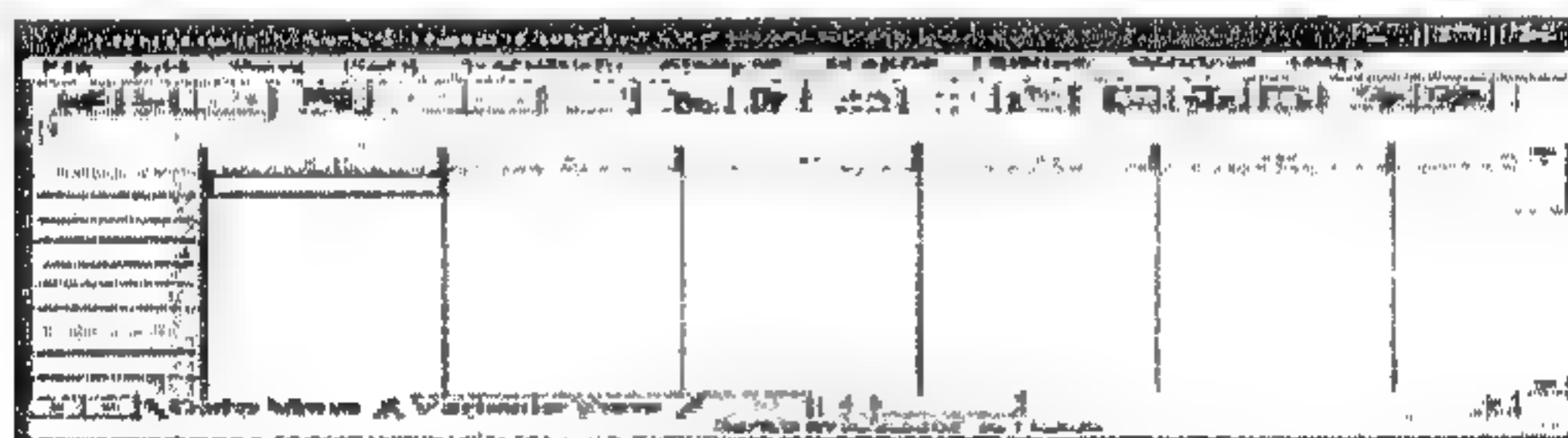
$$X = 5, 4, 2, 1$$

$$Y = 7, 6, 3, 4$$

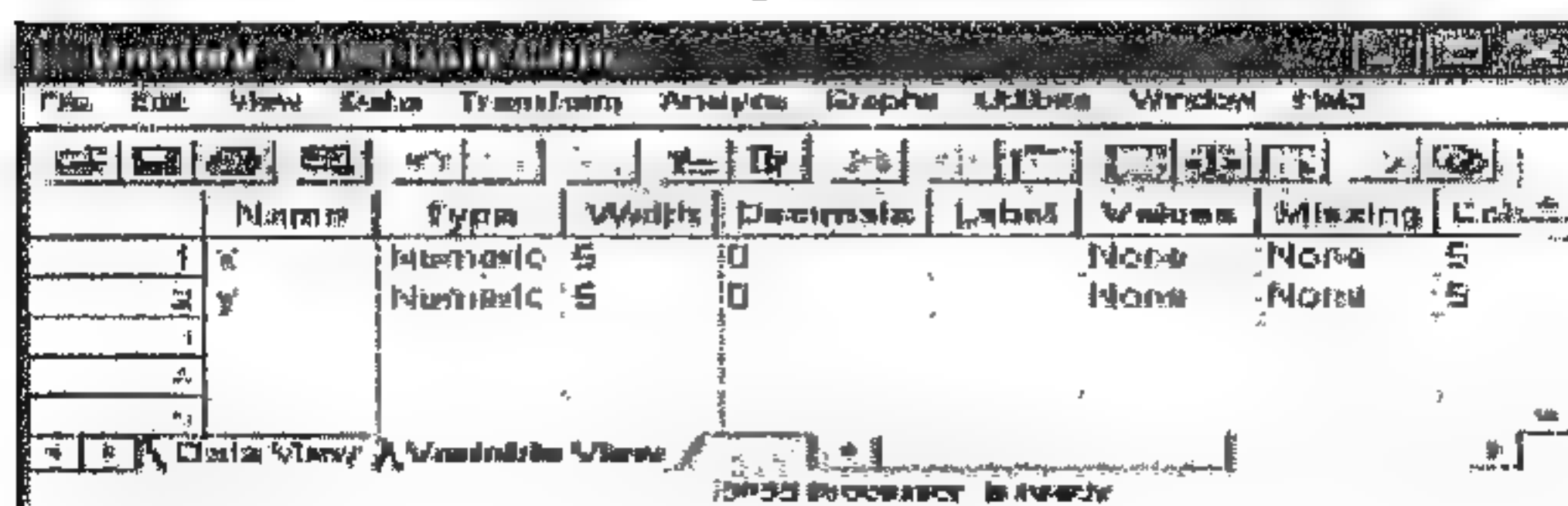
- 1- هل تعتقد بوجود علاقة بين العلامتين؟ ولماذا؟
- 2- احسب معامل ارتباط بيرسون بين هذين المتغيرين؟
- 3- احسب معامل ارتباط سيرمان للرتب بين هذين المتغيرين؟

Start-- Programs--SPSS for Windows--SPSS10.0 for Windows--Type in data-- Ok

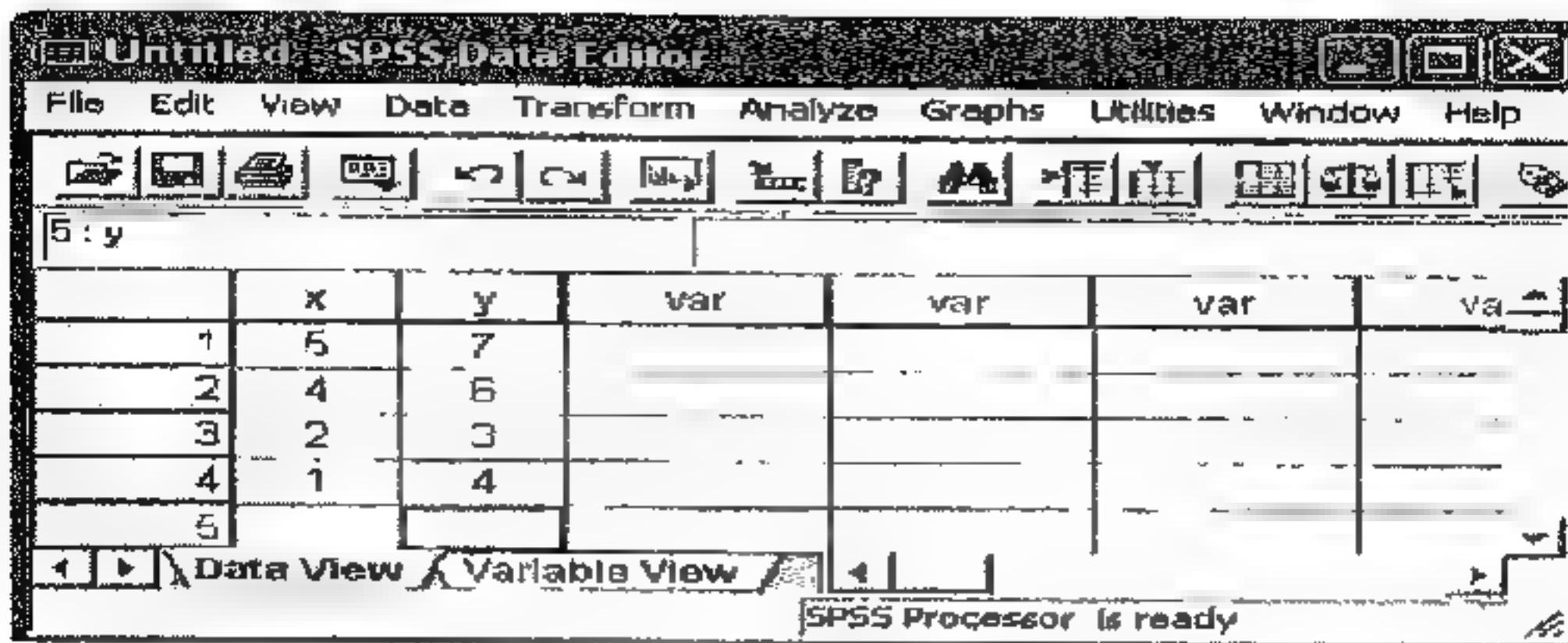
تظهر لديك الشاشة المبينة أدناه:



الخطوات المتبعة لتعريف المتغيرات: لنفرض أننا نريد تعريف المتغيرين X ، y انقر على Variable View الموجودة على شريط الحالة فتظهر الشاشة أدناه



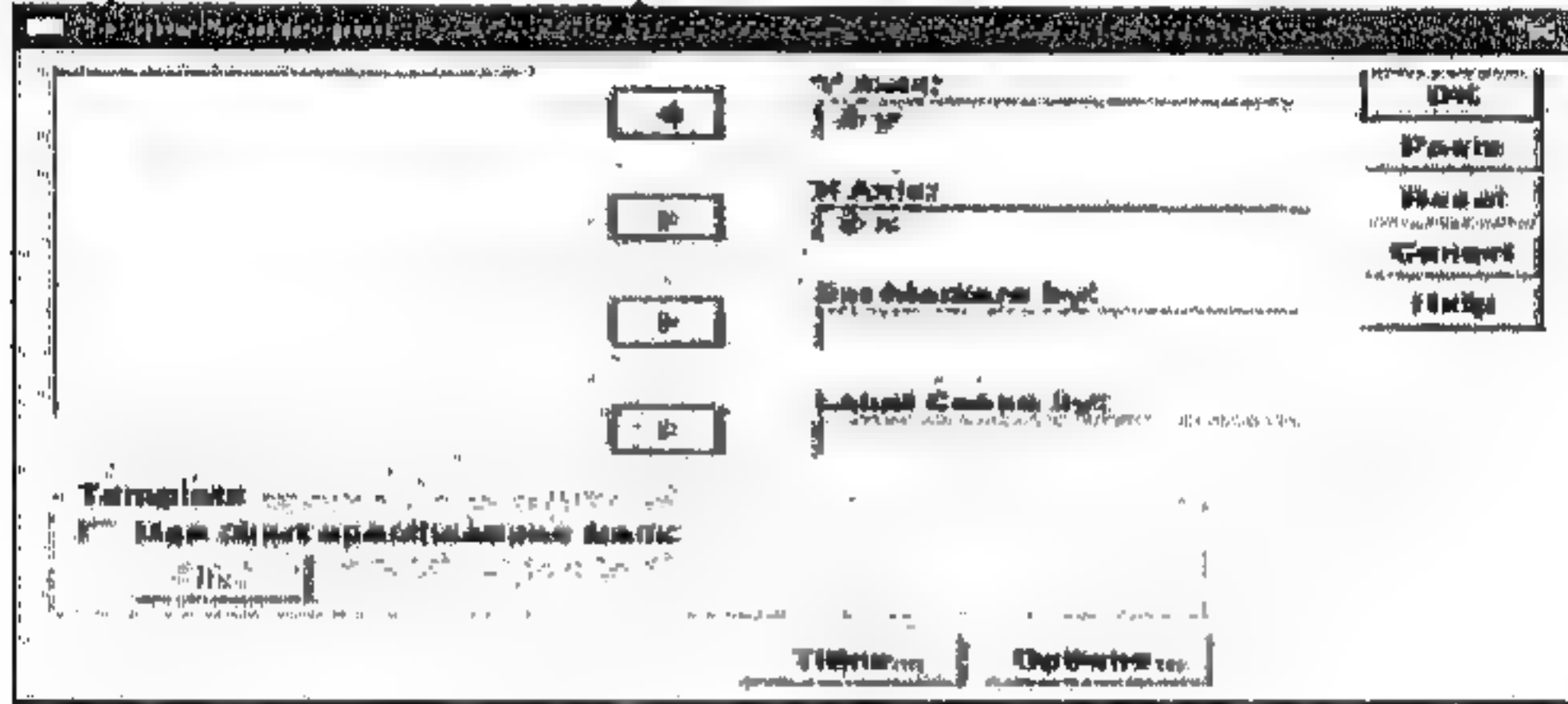
إدخال البيانات Data Input



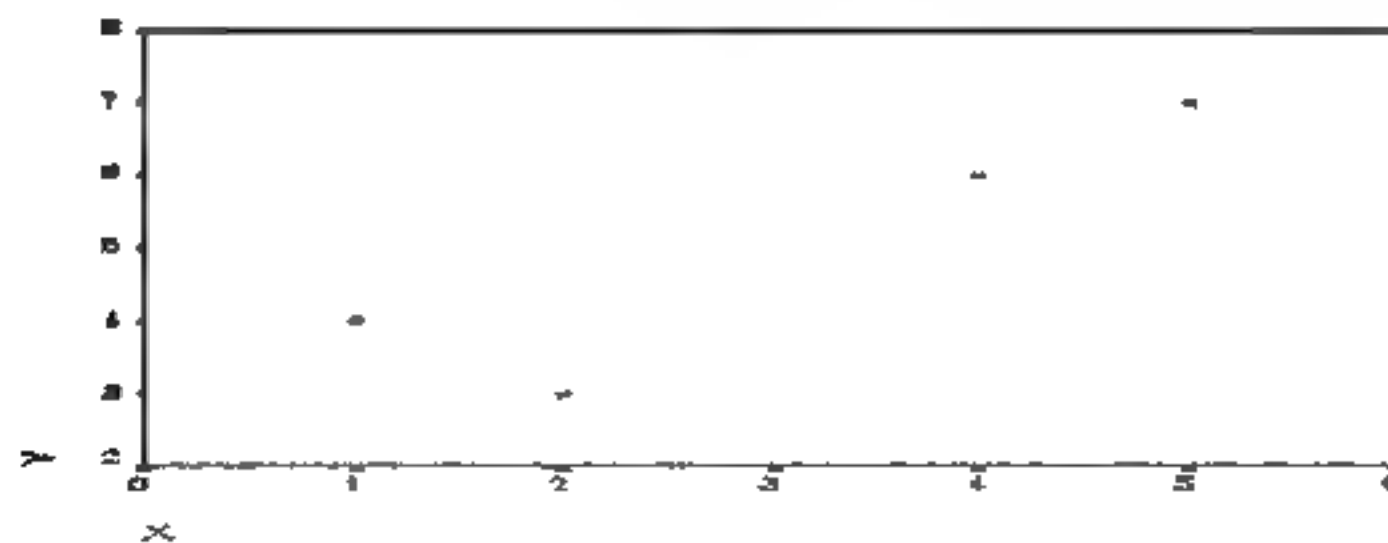
1- هل تعتقد بوجود علاقة بين العلامتين؟ ولماذا؟

للإجابة على هذا السؤال نقوم برسم شكل الانتشار

Graphs – Scatter... - Simple – Define



ثم نضغط زر OK فيظهر شكل الانتشار كما هو موضح أدناه:

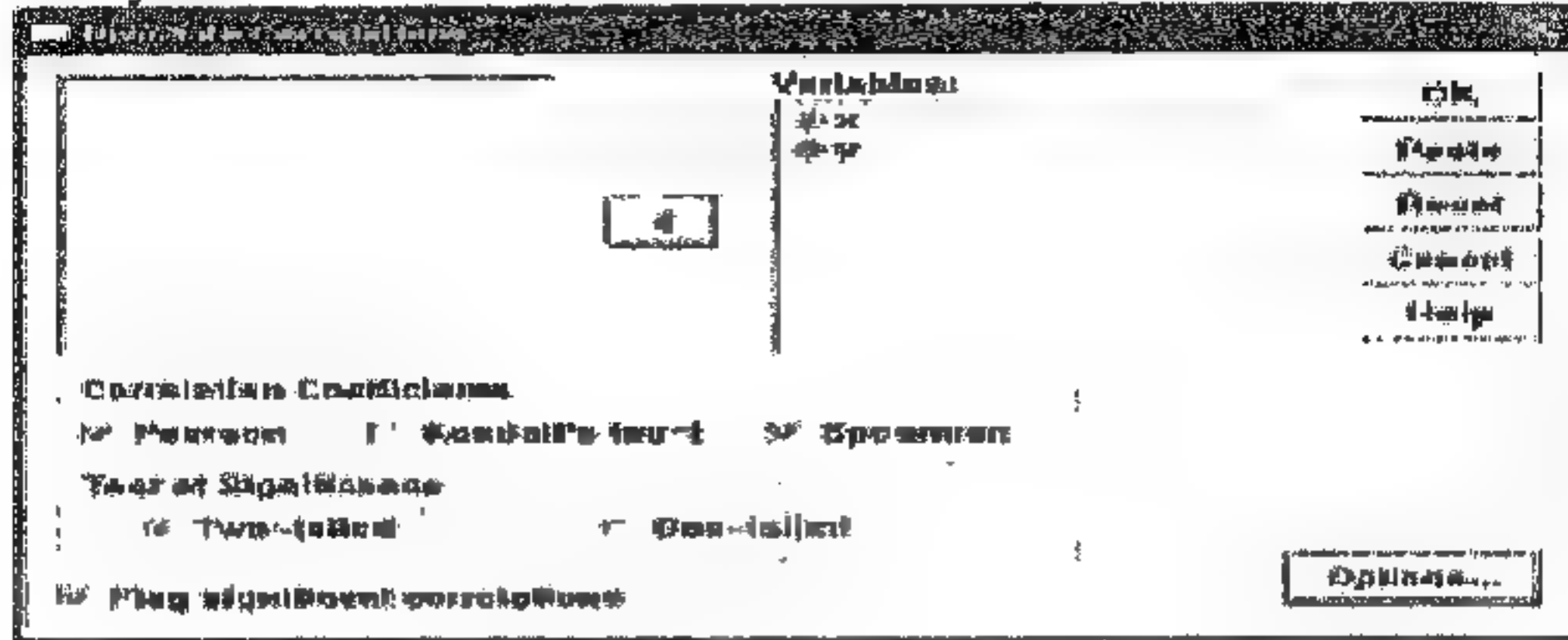


شكل الانتشار يدل على وجود علاقة طردية

2- احسب معامل ارتباط بيرسون بين هذين المتغيرين؟

3- احسب معامل ارتباط سبيرمان للرتب بين هذين المتغيرين؟

Analyze – Correlate - Bivariate



اختر المتغيرين الكميي x و y وانقلهما الى نافذة Variables:

من Correlation Coefficients اختر Person و Spearman بالنقر على مربعهما

انقر زر Ok تظهر شاشة المخرجات أدناه:

Correlations

			X	Y
X	Pearson Correlation		1.000	.800
	Sig. (2-tailed)		.	.100
	N		4	4
Y	Pearson Correlation		.800	1.000
	Sig. (2-tailed)		.100	.
	N		4	4

			X	Y
Spearman's rho	X	Correlation Coefficient	1.000	.800
		Sig. (2-tailed)		.200
		N	4	4
	Y	Correlation Coefficient	.800	1.000
		Sig. (2-tailed)	.200	.
		N	4	4

مثال: حساب معاملات الارتباط الجزئية Partial Correlation

استخدم مفهوم العلاقات مع الآخرين

المتغيرات: العلاقات مع الزملاء Friends، العلاقات مع الطلاب Students، العلاقات

مع المدراء Managers، العلاقات العامة General

سؤال الدراسة:

هل أعضاء هيئة التدريس الذين لديهم علاقات عالية في احد المجالات، يكون لديهم علاقات جيدة في الابعاد الأخرى اذا كان لديهم المستوى نفسه من العلاقات العامة.

* ادخال البيانات Input Data

Person	Person	Person	Person
1	12	12	12
2	14	14	14
3	16	16	16
4	18	18	18
5	20	20	20
6	22	22	22
7	24	24	24
8	26	26	26
9	28	28	28
10	30	30	30

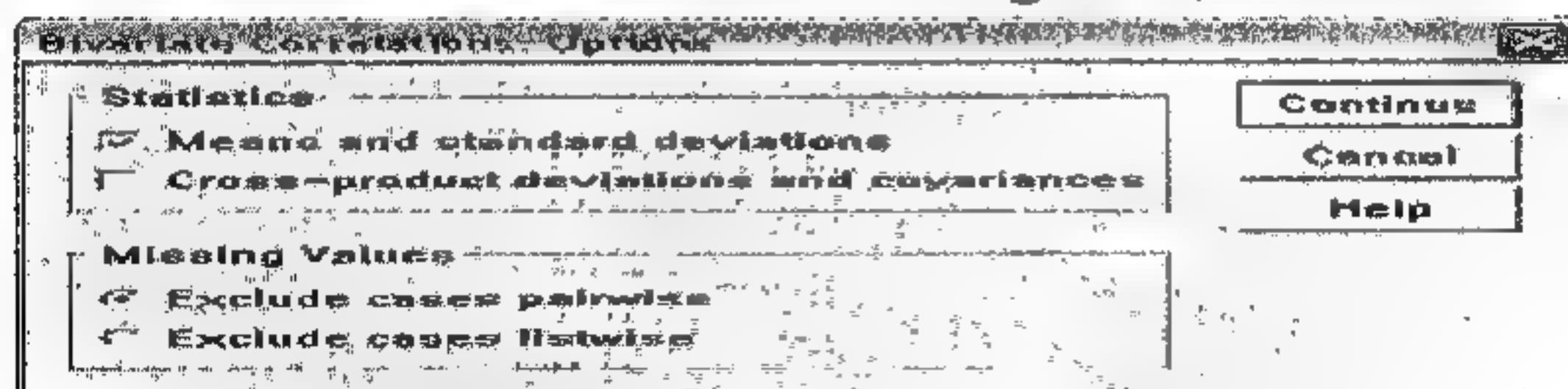
Analyze – Correlate – Bivariate...



Variables: انقل المتغيرات الى نافذة

Person انقر مربع الاختيار لمعامل ارتباط

Options... انقر زر Means and standard deviations انقر مربع اختيار



انقر زر Continue، ثم انقر زر Ok تظهر شاشة المخرجات أدناه:

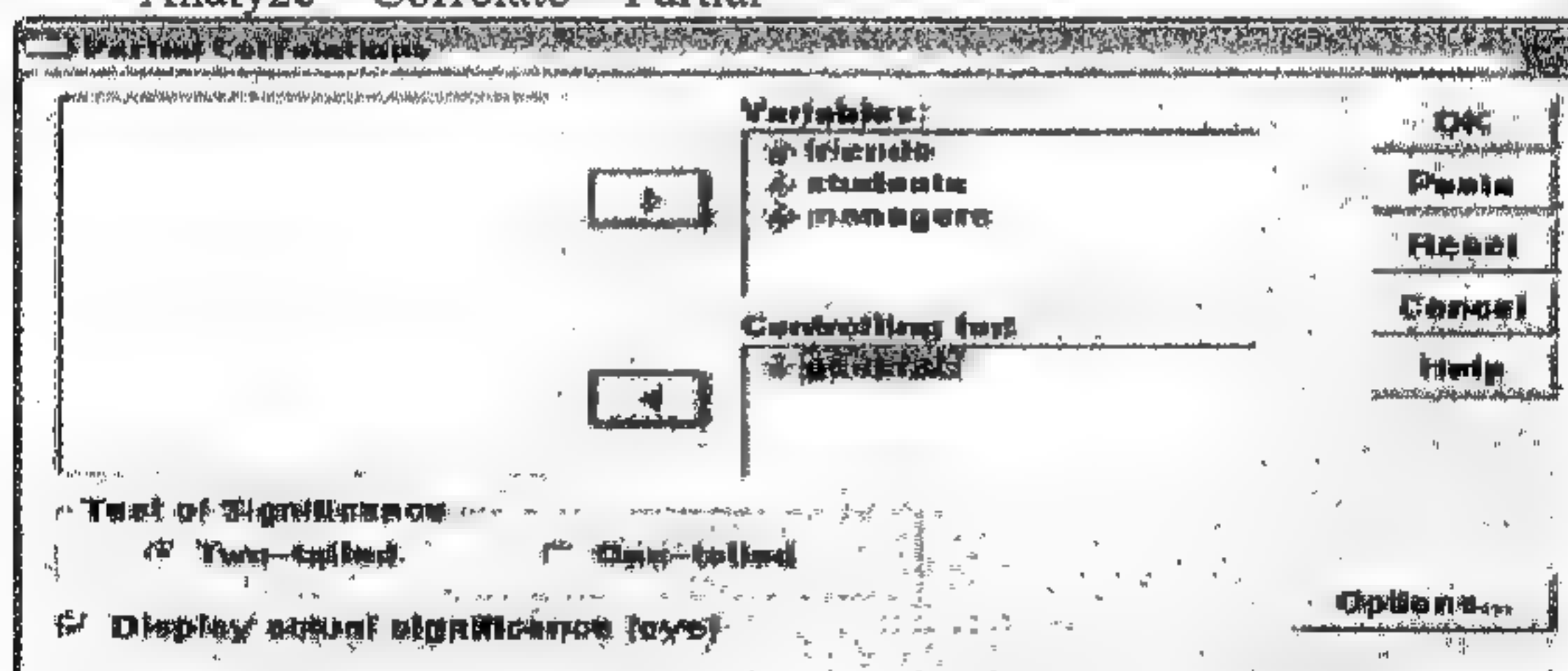
Correlations

Descriptive Statistics الإحصاءات الوصفية للمتغيرات التي تم اختيارها			
	Mean	Std. Deviation	N
FRIENDS	5.60	2.066	10
STUDENTS	6.10	1.853	10
MANAGERS	5.90	2.132	10
GENERAL	6.50	1.509	10

Correlations معاملات الارتباط بين المتغيرات التي تم اختيارها					
		FRIENDS	STUDENTS	MANAGERS	GENERAL
FRIENDS	Pearson Correlation	1	.912**	.924**	.820**
	Sig. (2-tailed)		.000	.000	.004
	N	10	10	10	10
STUDENTS	Pearson Correlation	.912**	1	.790**	.934**
	Sig. (2-tailed)	.000		.006	.000
	N	10	10	10	10
MANAGERS	Pearson Correlation	.924**	.790**	1	.873**
	Sig. (2-tailed)	.000	.008		.033
	N	10	10	10	10
GENERAL	Pearson Correlation	.820**	.934**	.873**	1
	Sig. (2-tailed)	.004	.000	.033	
	N	10	10	10	10

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

Analyze – Correlate – Partial



ضع المتغيرات friends, students, managers في نافذة: Variables وضع المتغير general في نافذة: Controlling for، انقر زر... Options تظهر الشاشة أدناه:

Partial Correlations: Options

Statistics

☒ Means and standard deviations

☒ Zero-order correlations

Missing Values

☒ Exclude cases listwise

☐ Exclude cases pairwise

Continue

Cancel

Help

انقر مربعي Means and standard deviations و Zero-order correlations

انقر زر Continue ، ثم انقر زر Ok تظهر شاشة المخرجات أدناه:

→ Partial Corr

Variable	Mean	Standard Dev	Cases
FRIENDS	5.6000	2.0656	10
STUDENTS	6.1000	1.8529	10
MANAGERS	5.9000	2.1318	10
GENERAL	6.5000	1.5092	10



المتوسطات Mean والانحرافات المعيارية Standard Dev وعدد الأفراد Cases

--- PARTIAL CORRELATION COEFFICIENTS ---

Zero Order Partial

	FRIENDS	STUDENTS	MANAGERS	GENERAL
FRIENDS	.00000 [.00] P= .	-.0116 [.01] P= .000	-.2833 [.01] P= .000	-.4108 [.01] P= .004
STUDENTS	-.0116 [.01] P= .000	1.0000 [.00] P= .	-.7204 [.01] P= .004	-.0137 [.01] P= .000
MANAGERS	-.2833 [.01] P= .000	-.7204 [.01] P= .000	1.0000 [.00] P= .	-.0714 [.01] P= .033
GENERAL	-.4108 [.01] P= .004	-.0137 [.01] P= .000	-.0714 [.01] P= .033	1.0000 [.00] P= .

(Coefficients < .05, P < 2-tailed Significance)

" . " is printed if a coefficient cannot be computed

معاملات الارتباط الثنائية Zero-Order Correlations

- -PARTIAL CORRELATION COEFFICIENTS - -

Controlling for.. GENERAL

	FRIENDS	STUDENTS	MANAGERS
FRIENDS	1.0000 (0) P= .	.7127 (7) P= .031	.8774 (7) P= .002
STUDENTS	.7127 (7) P= .031	1.0000 (0) P= .	.6107 (7) P= .081
MANAGERS	.8774 (7) P= .002	.6107 (7) P= .081	1.0000 (0) P= .

(Coefficient / (D.F.) / 2-tailed Significance)

" . " is printed if a coefficient cannot be computed

معاملات الارتباط الجزئية Partial Correlations

النتيجة:

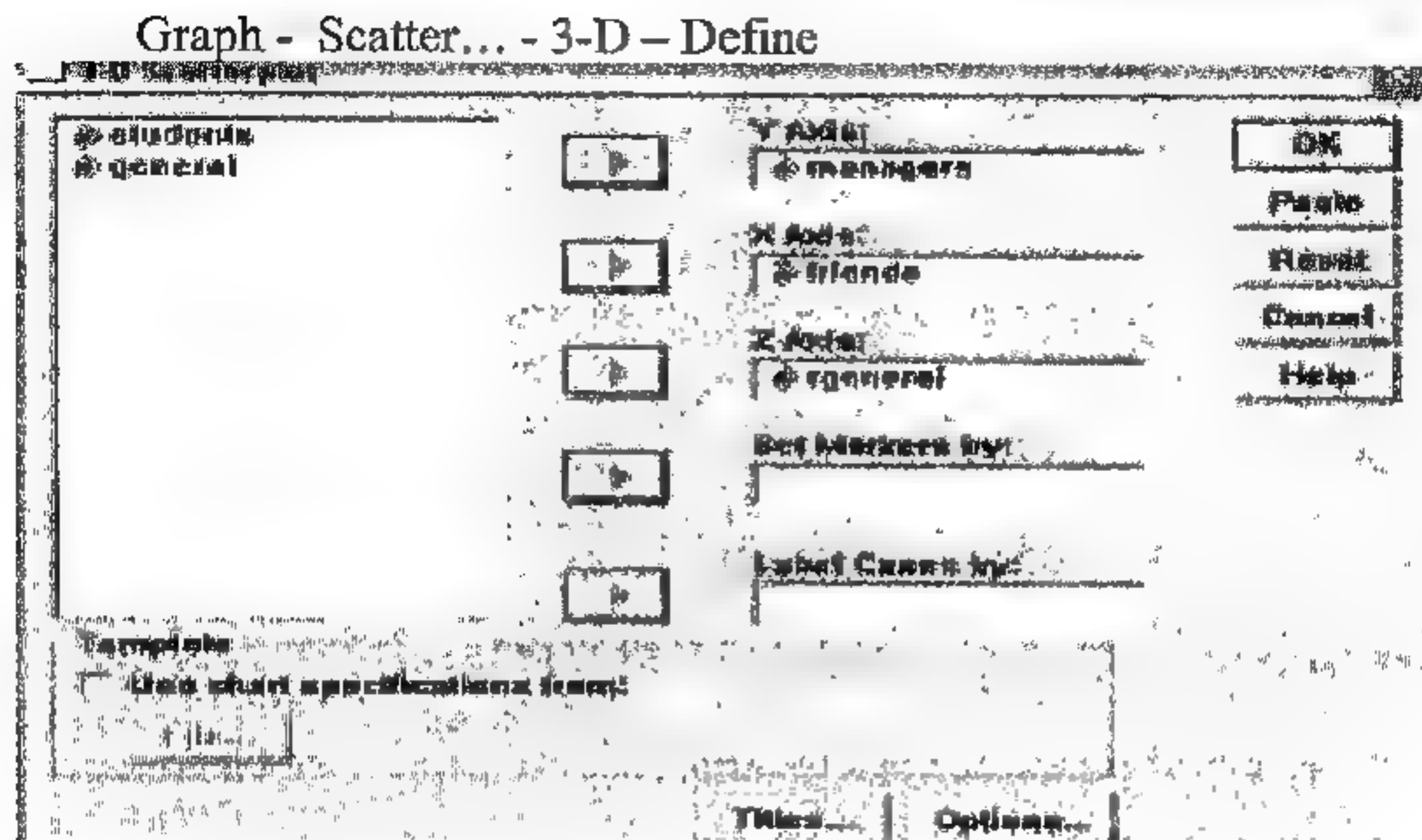
تم حساب المتوسطات Mean والانحرافات المعيارية Standard Dev وعدد الأفراد Cases لكل متغير من المتغيرات، كما حسبت معاملات الارتباط الثنائية Zero-Order Correlations فإذا كانت قيمة مستوى الدلالة اقل من المستوى المقبول 0.05 فإن معامل الارتباط يكون مقبولاً إحصائياً، وقد حسبت معاملات الارتباط الجزئية Partial Correlations فإذا كانت قيمة مستوى الدلالة اقل من المستوى المقبول 0.05 فإن معامل الارتباط الجزئي تكون مقبولة إحصائياً، أما إذا كانت قيمة مستوى الدلالة أكبر من المستوى المقبول 0.05 فإن قيمة معامل الارتباط الجزئي تكون غير مقبولة إحصائياً ويمكن إقرار عدم وجود علاقة بين المتغيرين.

وحتى نقلل من احتمال رفض الفرضية الصفرية وهي صحيحة (الخطأ من النوع الأول) يجب تعديل مستوى الدلالة ليصبح 0.05 مقسوماً على عدد معاملات الارتباط المحسوبة (3 في

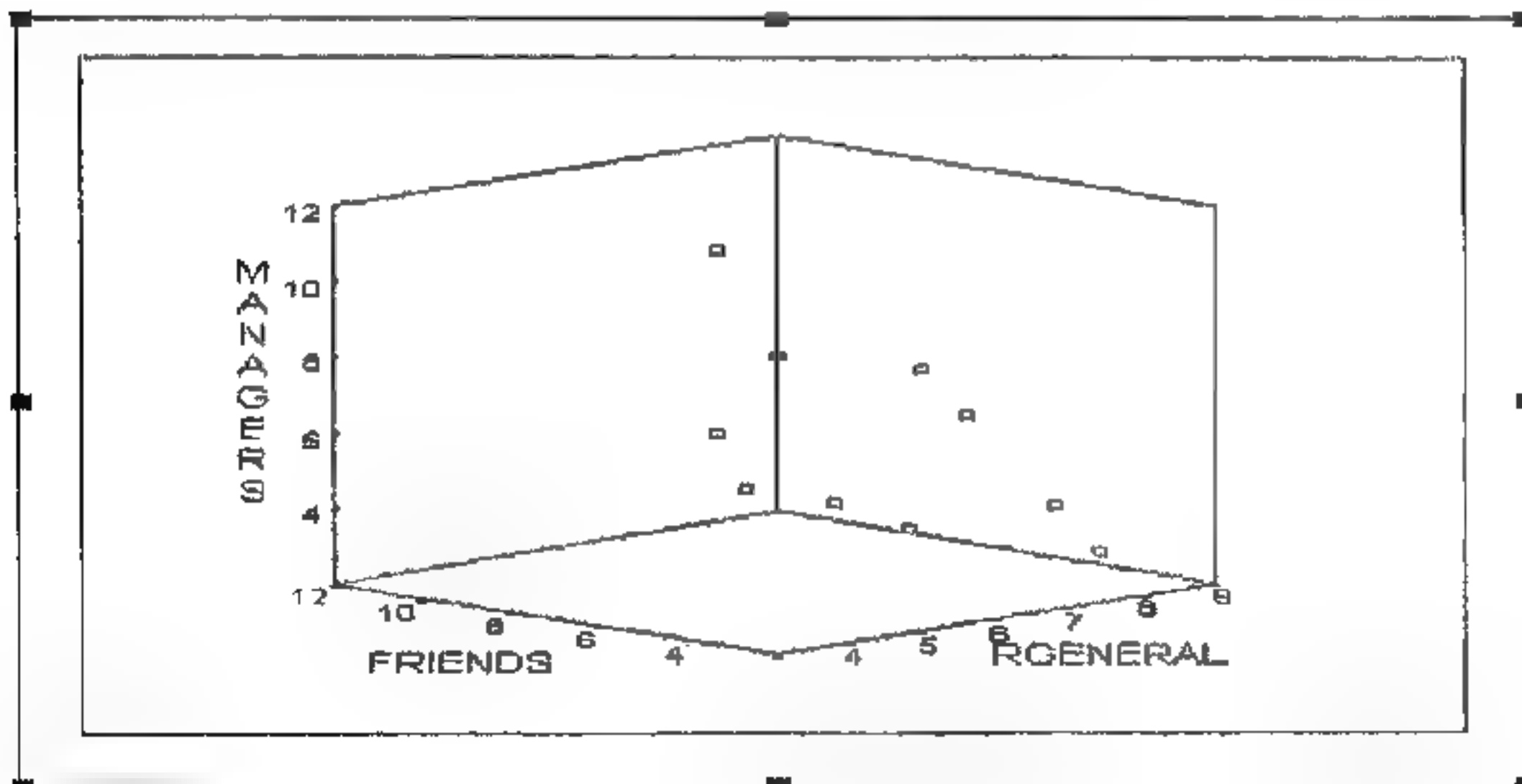
هذا المثال) لتصبح $0.0167 = 3/0.05$ ، وباستخدام هذا المعيار فإن معاملات الارتباط الجزئية بين Friends و Manager هو الارتباط الجزئي المقبول إحصائياً من أصل الارتباطات المحسوبة.

استخدام الرسوم البيانية لتوضيح النتائج

لوحة الانتشار ثلاثية الأبعاد 3D-Scatterplot



انقر زر Ok تظهر شاشة المخرجات أدناه:



لوحة الانتشار ثلاثية الأبعاد 3D-Scatterplot

6-7 تمارين Exercise

س1: ضع رمز الإجابة الصحيحة في الربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										
الرقم	11	12	13	14	15	16	17	18	19	20
الإجابة										

1- حسب معامل الارتباط بين X , Y فكان -0.65 ، عرف المتغيران $L = 2X - 0.2$ ،

$M = -3X + 0.3$ فإن معامل الارتباط الجديد بين L , M هو:

أ- (-0.35) ب- (-0.85) ج- (-0.65) د- 0.65

2- إذا كان $\Sigma(x-x') = 27$ ، وكان $\Sigma(x-x')^2 = 36$ ، وكان $\Sigma(y-y')^2 = 81$ ،

فإن معامل ارتباط بيرسون يساوي $r_{xy} = ?$

أ- (-27) ب- (-27) ج- (-27) د- (-27)
36 81 9 54

3- احدى القيم التالية يمثل معامل ارتباط سلبي عال جداً:

أ- (-0.7) ب- (0) ج- (-0.98) د- (-0.99)

4- إذا كانت قيم X كلها $5, 5, 5, 5, 5, \dots$ وقيم Y كلها $-3, -3, -3, -3, \dots$ فإن معامل ارتباط سبيرمان =

أ- 1 ب- (-1) ج- 0 د- 0.12

5- إذا أردنا قياس قوة العلاقة واتجاهها بين متغيرين دون البحث في العلاقة السببية فإن الأسلوب الإحصائي المناسب هو:

أ- تحليل الارتباط ب- تحليل الانحدار ج- التوقع الرياضي د- احتمال الحدث

6- إذا أعطيت بيانات عن متوسط درجات الحرارة وكمية الاستهلاك من الكهرباء خلال 50 يوم من أيام الصيف فإنه يمكن القول أن اتجاه العلاقة بين المتغيرين:

أ- عكسي ب- طردي ج- قوي د- ضعيف

- 7- نوع الارتباط بين مساحة المربع وأبعاده هو:
- أ- تام، بسيط ب- غير تام، بسيط ج- غير تام، متعدد د- تام، متعدد
- 8- عند حساب معامل الارتباط بين لون العيون ل (50) طالبة ونسبة الذكاء كان معامل الارتباط (0.9) فإنه يمكن تفسير هذا الارتباط القوي ل:
- أ- التأثير بمتغير ثالث فقط ب- التأثير بأكثر من متغير ج- العلاقة التبادلية د- عامل الصدفة
- 9- القيمة التي تمثل أقوى معامل ارتباط عكسي مما يلي هو:
- أ- (-2) ب- (-0.7) ج- (-0.90) د- (-1.5)
- 10- قام طالب بحساب معامل ارتباط بيرسون بين متغيرين فوجد أنه = (1.3) فإن ذلك يدل على:
- أ- خطأ في الحساب ب- عدم وجود ارتباط
ج- طردي تام د- عكسي تام
- 11- إذا كان مجموع مربعات فروق الرتب بين (6 قيم) للمتغيرين X, Y هو (50)، فإن معامل ارتباط سبيرمان للرتب بين X, Y يساوي:
- أ- (-1.43) ب- (-0.43) ج- (-0.57) د- (-0.43)
- 12- إذا كان $\Sigma(x-x') = 5$ ، وكان $\Sigma(x-x')^2 = 9$ ، وكان $\Sigma(y-y')^2 = 4$ ، فإن معامل ارتباط بيرسون يساوي $r_{xy} = ?$
- أ- (-5) ب- (-5) ج- (-5) د- (-5)
36 4 6 9
- 13- حسب معامل ارتباط سبيرمان بين X, Y فكان 0.6 فإذا كانت $n=10$ ، فإن Σd^2 يساوي
- أ- 44 ب- 66 ج- 0.4 د- 0.66
- 14- إذا كانت $r_1=0.6$ ، $r_2=0$ ، $r_3=0.43$ ، $r_4=-0.9$ فإن معامل الارتباط الذي يعبر عن أقوى علاقة هو:
- أ- r_1 ب- r_2 ج- r_3 د- r_4

س2: البيانات في الجدول أدناه هي لعينة من نزلاء الفندق حجمها $n=10$ تخص فترة إقامة كل نزيل في الفندق بالأيام X_i ومعدل الأجور في اليوم بالدينار Y والمطلوب:

رقم النزيل	الإقامة بالأيام X	تكلفة اليوم بالدينار Y
1	100	6
2	80	10
3	30	15
4	10	20
5	90	12
6	60	11
7	20	30
8	95	5
9	15	25
10	50	8

1. رسم شكل الانتشار بين X, Y

2. إيجاد معامل الارتباط البسيط r بين فترة الإقامة X_i ومعدل الأجور بالدينار Y

س3: الجدول التالي يضم بيانات تخص أعمار عينة من الموظفين X_1 والخبرة الوظيفية بالسنين X_2 والراتب الشهري بالدينار Y والمطلوب:

مسلسل	العمر بالسنوات X_1	سنوات الخبرة X_2	الراتب الشهري Y
1	25	2	150
2	30	4	170
3	32	5	190
4	35	7	200
5	37	8	220
6	40	10	250
7	45	12	300
8	50	15	350
9	55	18	370
10	60	20	500

1. إيجاد معامل الارتباط البسيط r بين سنوات الخبرة X_2 والراتب الشهري Y
2. إيجاد معامل الارتباط المتعدد R
3. إيجاد معامل ارتباط الجزئي $r_{y2.1}$ و $r_{21.y}$ مع تفسير النتائج.

س4: اجري بحث عن العلاقة بين طول لاعب كرة السلة والقدرة على التسجيل لدى عينة من اللاعبين، والباحث يفترض أن العلاقة هي نتيجة اللياقة البدنية، بمعنى أن اللاعبين يتدربون أكثر تصبح لديهم قدرة أكبر، ولفحص الفرضية تم تدوين بين طول اللاعب وقدرته على التسجيل لدى 10 لاعبين، كما قام بتدوين عدد الساعات الأسبوعية التي يستغرقها اللاعب في التمرين، إذا أراد الباحث فحص العلاقة بين طول اللاعب والقدرة على التسجيل بعد استبعاد أثر عدد ساعات التمرين (بافتراض أن جميع اللاعبين يتدربون العدد نفسه من الساعات).

طول اللاعب	القدرة على التسجيل	عدد ساعات التمرين
180	20	20
185	18	20
190	22	20
200	17	20
188	15	20
189	20	20
201	25	20
195	16	20
197	20	20
202	22	20

الفصل الثامن

اختبار الفرضيات حول النسب Hypothesis Testing Regarding Proportions

- 1-8 اختبار الفرضيات حول النسب
- 2-8 اختبار الفرضيات حول نسبة واحدة
- 3-8 اختبار الفرضيات حول نسبتين مستقلتين
- 4-8 اختبار الفرضيات حول نسبتين للبيانات الغير مستقلة
- 5-8 استخدام برمجية SPSS في حل المسائل.
- 6-8 تمارين Exercise

الفصل الثامن

اختبار الفرضيات حول النسب

Hypothesis Testing Regarding Proportions

1-8 اختبار الفرضيات حول النسب

Hypothesis Testing Regarding Proportions

يستخدم لإصدار تعميمات تتعلق بالنسب، وتستخدم للظواهر التي لا يمكن قياسها ولكن يمكن تقدير نسبة وقوعها. وهنا لا بد ان نتذكر أن المتغيرات التي نتعامل معها هي متغيرات تصنيفية يعني أنها تقع في المستوى الاسمي Nominal من مستويات القياس.

مثال: هل تختلف نسبة الموافقين على البحث التربوي باختلاف الجنس؟

مثال: هل تختلف نسبة المعاقين في الأردن عن نسبتهم في سوريا؟

في حالة المجتمع:

يرمز لنسبة وقوع الظاهرة في المجتمع بالرمز P وهي نسبة النجاح.

ويرمز لنسبة عدم وقوع الظاهرة في المجتمع بـ Q وهي نسبة الفشل، وهي $(1-P)$.

أما في حالة العينة:

يرمز لنسبة وقوع الظاهرة في العينة بالرمز p' وهي نسبة النجاح.

ويرمز لنسبة عدم وقوع الظاهرة في العينة بالرمز P وهي نسبة الفشل.

2-8 اختبار الفرضيات حول نسبة واحدة

Single Proportion (P)

• إذا كانت الفرضية غير متجهه

نسبة الحاصين على شهادات عليا في المجتمع الأردني.

نسبة وحدات الإنتاج الصالحة في مصنع معين.

يؤول هذا التوزيع إلى التوزيع الطبيعي عندما تصبح $n \geq 5$ وكذلك $n(n-1) \geq 0$

ويكون الخطأ المعياري للتوزيع $\sigma_L = \sqrt{n(1-n)}$

$$Z = \frac{p' - P}{\sqrt{PQ/n}}$$

حيث أن: n : حجم العينة

نسبة وقوع الظاهرة في العينة ب p' وهي نسبة النجاح.

نسبة وقوع الظاهرة في المجتمع ب P وهي نسبة النجاح.

نسبة عدم وقوع الظاهرة في المجتمع ب Q وهي نسبة الفشل، وهي $(1-P)$.

مثال 8-2: يدعي احد المصانع بأن ما لا يقل عن 94% من إنتاجه مطابق لمواصفات المطلوبة، وللتأكد من صحة هذا الادعاء أخذت عينة من إنتاج المصنع حجمها 100 وحدة وبعد فحصها وجد بأن 90% منها كانت مطابقة للمواصفات. فهل تدعم هذه النتيجة ادعاء المصنع بمستوى دلالة = 0.05

$$H_0: P \geq 0.94$$

$$H_1: P < 0.94$$

$$Z_{0.05} = 1.64 \text{ القيمة الحرجة}$$

$$Z = \frac{p' - P}{\sqrt{PQ/n}} = \frac{0.90 - 0.94}{\sqrt{(0.94)(0.06)/100}} = \frac{-0.04}{0.02234} = -1.7899 = -1.79$$

النتيجة: بما أن القيمة Z المحسوبة (-1.79) أقل من القيمة الحرجة 1.64 نقبل الفرضية الصفرية ونستنتج بأن ادعاء صاحب المصنع صحيح، وأنه ليس هناك فرق جوهري بين نسبة المجتمع P ونسبة العينة p' عند مستوى الدلالة $\alpha=0.05$

مثال 8-2: يتوفر في الأسواق دواء وعلى أساس نسبة نجاحه في تخفيض توتر الأعصاب هي 60% وظهر دواء جديد لنفس المرض كان قد تم تجربته على عينة تتكون من 100 شخص ودلت النتائج على شفاء 70 شخصاً منهم باستخدام هذا الدواء الجديد. فهل يمكن الاستنتاج من أن الدواء الجديد هو أفضل من النوع المتوفر في الأسواق عند مستوى معنوية 0.05

الحل:

$$H_0: P = 0.6$$

$$H_1: P > 0.6$$

قيمة Z الجدولية عند مستوى الدلالة $\alpha=0.05$ الاختبار بذييل واحد، $Z_{0.05} = 1.64$

قيمة Z المحسوبة:

$$Z = \frac{p' - P}{\sqrt{PQ/n}} = \frac{0.7 - 0.6}{\sqrt{(0.6)(0.4)/100}} = \frac{0.1}{0.0489} = 2.044 = 2.04$$

النتيجة: بما أن القيمة Z المحسوبة (2.04) أكبر من القيمة الحرجة ($Z_{0.05} = 1.64$) نرفض الفرضية الصفرية H_0 القائلة أن نسبة نجاح الدواء الجديد هي 60% ونستنتج بأن نسبة نجاحه أكثر من 60% أي أنه أفضل من الدواء المستخدم في الأسواق.

• إذا كانت الفرضية متجهه

$$Z = \frac{L - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}}$$

مثال 8-3: إذا كان من المتوقع أن تكون نسبة الإناث في الدولة المتقدمة من بين ذوي الأعمار التي تزيد عن 50 سنة تساوي 0.5 ورجب باحث في معرفة إن كانت هذه النسبة تزيد عن ذلك في الأردن. فاختار عينة عشوائية طبقية حسب المستوى الاقتصادي الاجتماعي ممن تزيد أعمارهم عن 50 سنة. وكان حجم العينة 400 شخص فوجد أن عدد الإناث في هذه العينة 242 فهل تدعم هذه البيانات ما يدعيه الباحث؟

$$H_0: \pi = 0.5$$

$$H_1: \pi > 0.5$$

$$\alpha = 0.05 \quad Z = 1.64 \quad \text{القيمة الحرجة}$$

$$\alpha = 0.01 \quad Z = 2.34 \quad \text{القيمة الحرجة}$$

$$L = 242/400 = 0.605 \quad \text{النسبة}$$

القيمة المحسوبة

$$Z = \frac{L - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}}$$

$$Z = \frac{0.605 - 0.5}{\sqrt{\frac{0.5(1-0.5)}{400}}}$$

$$Z = \frac{0.105}{\sqrt{\frac{0.25}{400}}} = \frac{0.105}{0.025} = 4.2$$

بمسا أن قيمة Z المحسوبة (4.2) < الحرجة (1.64، 2.34) على مستوى $\alpha = 0.01$ يمكن أن نقول أن نسبة الإناث في الأردن من بين الذين تتجاوز أعمارهن 50 سنة تزيد بدلالة إحصائية مرتفعة ($P < 0.01$) عن النسبة المناضرة لها في الدول المتقدمة.

3-8 اختبار الفرضيات حول نسبتي مستقلتين

تكون النسبتان مستقلتين إذا كانتا تعودان إلى مجتمعين مستقلين.

مثال:

مقارنة نسبة المقبولين من الذكور لوظيفة معينة إلى نسبة المقبولات من الإناث.

مقارنة نسبة المصابين بالسرطان من الذكور إلى نسبة المصابات من الإناث.

مقارنة نسبة المقبولين من الذكور في الجامعة إلى نسبة المقبولات من الإناث.

$$H_0: \pi_1 = \pi_2$$

π_1 : النسبة في المجتمع الأول. π_2 : النسبة في المجتمع الثاني.

σ^2_L : التقدير للخطأ المعياري لتوزيع المعاينة.

$$\sigma^2_L = \pi(1-\pi) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)$$

$$\pi^{\wedge} = \frac{F1 + F2}{n1 + n2}$$

F1 : التكرار للظاهرة في العينة الأولى. F2 : التكرار للظاهرة في العينة الثانية.

n1 : عدد الأفراد في العينة الأولى. n2 : عدد الأفراد في العينة الثانية.

$$Z = \frac{L1 - L2}{\sqrt{\pi^{\wedge} (1 - \pi^{\wedge}) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

مثال 4-8: جرى تطبيق طريقتين لتعليم مفهوم التسارع على مجموعتين عدد الأفراد في

كل منها 100 وفي نهاية التدريس وجد بأن عدد الذين أتقنوا المفهوم في المجموعة الأولى = 68

وأن العدد المناظر في المجموعة الثانية = 54 ، فهل تختلف الطريقتان جوهرياً في نسب الذين

يتقنون المفهوم في كل منها؟

$$H_0: \pi_1 = \pi_2$$

$$H_1: \pi_1 \neq \pi_2$$

القيمة الحرجة على مستوى $\alpha=0.05$ هي 1.69 وعلى مستوى $\alpha=0.01$ هي 2.58
القيمة المحسوبة:

$$L1 = 68/100 = 0.68$$

$$L2 = 54/100 = 0.54$$

$$\pi^{\wedge} = \frac{F1 + F2}{n1 + n2} = \frac{68 + 54}{100 + 100} = 0.61$$

$$Z = \frac{L1 - L2}{\sqrt{\pi^{\wedge} (1 - \pi^{\wedge}) \left(\frac{1}{n1} + \frac{1}{n2} \right)}}$$

$$Z = \frac{0.68 - 0.54}{\sqrt{0.61 (1 - 0.61) \left(\frac{1}{68} + \frac{1}{54} \right)}} = 2.03$$

النتيجة: بما أن Z المحسوبة (2.03) تقع في منطقة الرفض للفرضية الصفرية على مستوى $\alpha=0.05$ يمكن القول أن نسبي متقني مفهوم التسارع من المجموعتين تختلفان اختلافاً ذا دلالة إحصائية ($P < 0.05$).

P_1 : نسبة النجاح في المجتمع الأول. P_2 : نسبة النجاح في المجتمع الثاني.

p'_1 : نسبة النجاح في العينة الأول. p'_2 : نسبة النجاح في العينة الثانية.

$$\mu_{p'_1} = P_1$$

$$\sigma_{p'_1} = \sqrt{\frac{P_1 Q_1}{n1}}$$

مثال 5-8: في استطلاع لآراء 100 مدرس و 100 مدرسة حول تنفيذ برنامج

تربوي معين كان عدد الموافقين من المدرسين 33 ومن المدرسات 18، هل تختلف النسبة في

المدرسين عما هي عليه في المدرسات؟ وذلك باستخدام اختبار Z

أ- تختلف وذلك لأن $Z=1.177$ وهي ذات دلالة إحصائية على مستوى $\alpha=0.05$

ب- تختلف وذلك لأن $\chi^2_2=5.92$ وهي ذات دلالة إحصائية على مستوى $\alpha=0.05$

ج- لا تختلف وذلك لأن $Z=1.177$ وهي ذات دلالة إحصائية على مستوى $\alpha=0.05$

د- لا تختلف وذلك لأن $\chi^2_2=5.92$ وهي ذات دلالة إحصائية على مستوى $\alpha=0.05$

$$H_0: \pi_1 = \pi_2$$

$$H_1: \pi_1 \neq \pi_2$$

القيمة الحرجة على مستوى $\alpha=0.05$ هي 1.69 وعلى مستوى $\alpha=0.01$ هي 2.58

القيمة المحسوبة:

$$L1 = 33/100 = 0.33$$

$$L2 = 18/100 = 0.18$$

$$\pi^{\wedge} = \frac{F1 + F2}{n1 + n2} = \frac{33 + 18}{100 + 100} = 0.255$$

$$Z = \frac{L1 - L2}{\sqrt{\pi^{\wedge} (1 - \pi^{\wedge}) \left(\frac{1}{n1} + \frac{1}{n2} \right)}}$$

$$Z = \frac{0.33 - 0.18}{\sqrt{0.255 (1 - 0.255) \left(\frac{1}{33} + \frac{1}{18} \right)}} = 1.177$$

النتيجة:

بما أن Z المحسوبة (1.177) تقع في منطقة القبول للفرضية الصفرية على مستوى $\alpha=0.05$ يمكن القول أن نسبي متقني مفهوم التسارع من المجموعتين لا تختلفان اختلافاً ذا دلالة إحصائية ($P<0.05$).

4-8 اختبار الفرضيات حول نسبتي للبيانات الغير مستقلة (المتراصة)

تكون النسبتان غير مستقلتان إذا كانتا تعودان إلى مجتمعين غير مستقلين.

مثال: نسبة الموافقين على التعليم المختلط في الجامعات قبل استماعهم لمحاضرة تبين رأي الدين في ذلك وبعد استماعهم للمحاضرة.

$$H_0: \pi_1 = \pi_2$$

يستخدم اختبار χ^2 حيث تفرغ الإجابات في جدول التصاحب Contingency table حيث ترصد التكرارات داخل الجدول.

بعد			
موافق	معارض	نوع القرار	
b	a	موافق	قبل
d	c	معارض	

$$\chi_1^2 = \frac{(d - a)^2}{d + a}$$

القيمة المحسوبة

نستخرج القيمة الحرجة للرفض أو القبول باستخدام جدول χ^2 بدرجة حرية واحدة وحسب مستوى الدلالة المطلوب.

ترفض الفرضية الصفرية إذا كانت قيمة χ^2 المحسوبة $>$ أكبر من القيمة الحرجة.
تقبل الفرضية الصفرية إذا كانت قيمة χ^2 المحسوبة $<$ أقل من القيمة الحرجة.

مثال 5-8: تم استطلاع آراء 60 من التربويين حول مسألتين الأولى تتعلق بتأنيث التعليم في الصفوف الابتدائية الثلاثة الأولى والثانية تتعلق بتأنيث التعليم في المرحلة الابتدائية بأكملها. فإذا وافق 42 منهم على المسألة الأولى ووافق 34 منهم على المسألة الثانية. وكان جميع الذين وافقوا على المسألة الثانية قد وافقوا على المسألة الأولى. فهل تدل هذه البيانات على وجود اختلاف جوهري في نسب الموافقين على المسألتين؟

$$H_0: \pi_1 = \pi_2$$

$$H_1: \pi_1 \neq \pi_2$$

القيم الحرجة للمختبر الإحصائي χ^2 بدرجة حرية واحدة وحسب مستوى الدلالة

$$\chi^2_{(1,0.95)} = 3.84 \text{ هي } \alpha=0.05, \text{ وعلى مستوى } \alpha=0.01 \text{ هي } \chi^2_{(1,0.99)} = 6.64$$

القيمة المحسوبة:

	المسألة الثانية			
	موافق	معارض	نوع القرار	
42	34	8	موافق	المسألة الأولى
18	0	18	معارض	
60	34	26		

$$\chi^2 = \frac{(d-a)^2}{d+a} = \frac{(0-8)^2}{0+8} = \frac{64}{8} = 8$$

النتيجة:

بما أن القيمة المحسوبة ($\chi^2 = 8$) أكبر من القيمة الحرجة ($\chi^2_{(1,0.99)} = 6.64$) وعلى مستوى $\alpha=0.01$ نستطيع القول بأن نسبة التربويين الموافقين على تأنيث التعليم في الصفوف

السلسلة الأولى تختلف جوهرياً ($P < 0.01$) عن نسبة الموافقين على مثل هذا النوع من التعيم وذلك عندما يتم استطلاع آرائهم على المسألتين في نفس الوقت.

5-8 اختبار χ^2 لحسن المطابقة

The Chi-Square Goodness of fit

يستخدم للمقارنة بين توزيع العينة والتوزيع النظري للمجتمع الذي سحبت منه. اختبار كاي تربيع لحسن المطابقة، والغرض منه تحديد ما إذا كانت النسب الملاحظة تختلف جوهرياً عن قيم متوقعة أو نظرية لها . وهو مكافئ لاختبار Z للنسب.

$$\chi^2 = \sum (O_i - E_i)^2 / E$$

حيث O_i : التكرارات المشاهدة E_i : التكرارات المتوقعة

مثال: أجرى معلم دراسة لمعرفة نسب الرسوب في المواد التي يدرسها والتي على أساسها يتم تقسيم الطلاب إلى مستويات وقد قام بدراسة شملت 400 طالب كما يلي:

المستويات	1	2	3	4	5
عدد الطلاب	900	700	800	1100	700
نسبة الرسوب	%20	%15	%10	%30	%25

اختبر χ^2 لمعرفة مدى تطابق التوزيع.

الحل: نحسب التكرارات المتوقعة كما يلي

المستوى	التكرارات المشاهدة	التكرار المتوقع
1	900	$800 = 0.20 \times 4000$
2	700	$600 = 0.15 \times 4000$
3	800	$400 = 0.10 \times 4000$
4	1100	$1200 = 0.30 \times 4000$
5	700	$1000 = 0.25 \times 4000$
المجموع	4000	3000

إيجاد قيمة χ^2 المحسوبة من المعادلة $\chi^2 = \sum (O_i - E_i)^2 / E$

$$\chi^2 = (900-800)^2/800 + (700-600)^2/600 + (600-400)^2/400 + (1100-1200)^2/1200 + (700-000)^2/1000$$

$$\chi^2 = 12.5 + 16.67 + 100 + 8.3 + 9 = 146.47$$

قيمة χ^2 الجدولية $\chi^2_{(4, 0.01)} = 13.28$

القرار:

بما ان قيمة χ^2 الجدولية اقل من المحسوبة نرفض الفرضية الصفرية ونقبل البديلة، وهذا يعني ان توزيع العينة يختلف عن التوزيع النظري للمجتمع الذي سحبت منه.

2- اختبار χ^2 للاستقلالية The Chi-Square Test of Independence

لمعرفة اذا كان متغيرين مستقلين عن بعضهما.

يستخدم لمعسرة ما إذا كانت الآراء حول قضية معينة تختلف أو ترتبط مع متغير اسمي آخر، وعند استخدام هذا النوع من الاختبار فإنه يتم إنشاء جدول التوافق.

6-8 استخدام برمجية SPSS في حل المسائل.

1- اختبار كاي تربيع لحسن المطابقة، والغرض منه تحديد ما إذا كان النسب الملاحظة تختلف جوهرياً عن قيم متوقعة أو نظرية لها .

خطوات استخراجها من SPSS :

● ادخل البيانات تصاعدياً

● من قائمة Analyze اختر Nonparametric Tests

● اختر Chi-Square

● انقل المتغير إلى خانة Variable List

● من خانة Expected Value أشر على مربع Value

● ادخل النسب حسب التكرارات في محور البيانات

● قم بوزن الحالات على أساس التكرارات من قائمة Weight Cases Data

مثال: إذا كانت نسبة الذكاء بين طلاب الثانوية العامة تتوزع بين ثلاث مستويات (2: 7: 1) على التوالي مرتفعي الذكاء ومتوسطي الذكاء ومحدودي الذكاء، وأراد باحث أن يختبر عينة ممثلة من مدرسة محددة لهذا المجتمع بحجم 400 طالب فوجد أن عدد عالي الذكاء = 50 ومتوسط الذكاء = 290 ومحدودي الذكاء = 60 فهل يمكن اعتبار هذه العينة ممثلة للمجتمع.

2- اختبار كاي تربيع للاستقلالية: يستخدم لمعرفة ما إذا كانت الآراء حول قضية معينة تختلف أو ترتبط مع متغير اسمي آخر، وعند استخدام هذا النوع من الاختبار فإنه يتم إنشاء جدول التوافق.

خطوات استخراج SPSS :

- من قائمة Analyze اختر Descriptive Statistics ثم Crosstabs
- أنقل المتغيرات المطلوبة إلى الصفوف والأعمدة
- أنقر مربع Statistics ثم اختر Chi - Square
- أنقر مربع Cells وهذا يمكنك من إجراء إحصائيات إضافية
- قم بوزن الحالات على أساس التكرارات من قائمة Weight Cases Data

مثال: قام مدير التعليم بأحد مناطق المملكة بسؤال 200 من التربويين حول الاستمرار في السدوام الصيفي أم لا ، وسأل ثلاث فئات (مشرفين، معلمين، ومدراء المدارس) وكانت الإجابات على النحو التالي (المشرفين 38 موافق و 12 معارض، المعلمين 26 موافق و 102 معارض، المدراء 16 موافق و 6 معارض) فهل تدل البيانات على اختلاف الرأي حول هذه المسألة باختلاف الوظيفة؟

7-8 تمارين Exercise

تمرين (1) : ضع رمز الإجابة الصحيحة في المربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										

1- إذا كانت $n=100$ فإن القيمة العظمى لخطأ المعاينة المعطى بالعلاقة $\sigma_L = \sqrt{\pi(1-\pi)}$ ستكون عند القيمة التالية للمعلم π :

- أ- 0.5 ب- 0.9 ج- صفر د- 0.1

2- أي من الرموز التالية يشير إلى القيمة الحرجة لاختبار χ^2 عندما يكون مستوى الدلالة $\alpha=0.01$ وتكون درجة الحرية $\gamma=2$

- أ- $\chi^2_{2,0.01}$ ب- $\chi^2_{0.99}$ ج- $\chi^2_{2,0.99}$ د- $\chi^2_{2,0.90}$

3- أي من التالي يكون هو الأكبر؟

- أ- $\chi^2_{1,0.01}$ ب- $\chi^2_{1,0.99}$ ج- $\chi^2_{1,0.99}$ د- $\chi^2_{1,0.90}$

4- أي من التالي يكون هو الأكبر؟

- أ- $\chi^2_{1,0.95}$ ب- $\chi^2_{2,0.95}$ ج- $\chi^2_{3,0.95}$ د- $\chi^2_{1,0.90}$

5- اختبار يتألف من 40 فقرة حصل طالب على 26 فقرة بشكل صحيح هل يختلف هذا

العدد من الفقرات عما هو متوقع بالتخمين؟ إذا كانت الفقرات من نوع صح/خطأ.

أ- لا، يختلف لأن $Z=0.897$ و $\chi^2 = 3.6$ وهي ليست ذات دلالة احصائية.

ب- نعم، يختلف لأن $Z=0.897$ و $\chi^2 = 3.6$ وهي ليست ذات دلالة احصائية.

ج- لا، يختلف لأن $Z=0.897$ و $\chi^2 = 3.6$ وهي ذات دلالة احصائية.

د- نعم، يختلف لأن $Z=0.897$ و $\chi^2 = 3.6$ وهي ذات دلالة احصائية.

6- اختبار يتألف من 40 فقرة حصل طالب على 26 فقرة بشكل صحيح هل يختلف هذا

العدد من الفقرات عما هو متوقع بالتخمين؟ إذا كان لكل فقرة 4 بدائل.

أ- لا، يختلف لأن $Z=0.897$ و $\chi^2 = 3.6$ وهي ليست ذات دلالة احصائية.

ب- نعم، يختلف لأن $Z=5.84$ و $\chi^2=34.13$ هي ذات دلالة إحصائية على مستوى $\alpha=0.001$

ج- لا، يختلف لأن $Z=0.897$ و $\chi^2=3.6$ وهي ذات دلالة إحصائية.

د- لا، يختلف لأن $Z=5.84$ و $\chi^2=34.13$ هي ذات دلالة إحصائية على مستوى $\alpha=0.001$

7- في استطلاع لآراء 100 مدرس و 100 مدرسة حول تنفيذ برنامج تربوي معين كان عدد الموافقين من المدرسين 33 ومن المدرسات 18، هل يختلف النسبة في المدرسين عما هي عليه في المدرسات؟ وذلك باستخدام اختبار Z

أ- يختلف وذلك لأن $Z=2.433$ وهي ذات دلالة إحصائية على مستوى $\alpha=0.05$

ب- يختلف وذلك لأن $\chi^2_2=5.92$ وهي ذات دلالة إحصائية على مستوى $\alpha=0.05$

ج- لا يختلف وذلك لأن $Z=2.433$ وهي ذات دلالة إحصائية على مستوى $\alpha=0.05$

د- لا يختلف وذلك لأن $\chi^2_2=5.92$ وهي ذات دلالة إحصائية على مستوى $\alpha=0.05$

8- في استطلاع لآراء 100 مدرس و 100 مدرسة حول تنفيذ برنامج تربوي معين كان عدد الموافقين من المدرسين 33 ومن المدرسات 18، هل يختلف النسبة في المدرسين عما هي عليه في المدرسات؟ وذلك باستخدام اختبار Z

أ- يختلف وذلك لأن $Z=2.433$ وهي ذات دلالة إحصائية على مستوى $\alpha=0.05$

ب- يختلف وذلك لأن $\chi^2_2=5.92$ وهي ذات دلالة إحصائية على مستوى $\alpha=0.05$

ج- لا يختلف وذلك لأن $Z=2.433$ وهي ذات دلالة إحصائية على مستوى $\alpha=0.05$

د- لا يختلف وذلك لأن $\chi^2_2=5.92$ وهي ذات دلالة إحصائية على مستوى $\alpha=0.05$

الفصل التاسع

تحليل التباين

Analysis of Variance

- | | |
|--|------------------------------|
| 1-9 مقدمة | Introduction |
| 2-9 تحليل التباين الأحادي | One-Way Analysis of Variance |
| 3-9 تحليل التباين الثنائي | Tow-Way Analysis of Variance |
| 4-9 تحليل التغاير | Analysis of Covariance |
| 5-9 استخدام برمجية SPSS في حل المسائل. | |
| 6-9 تمارين | Exercise |

الفصل التاسع

تحليل التباين

Analysis of Variance

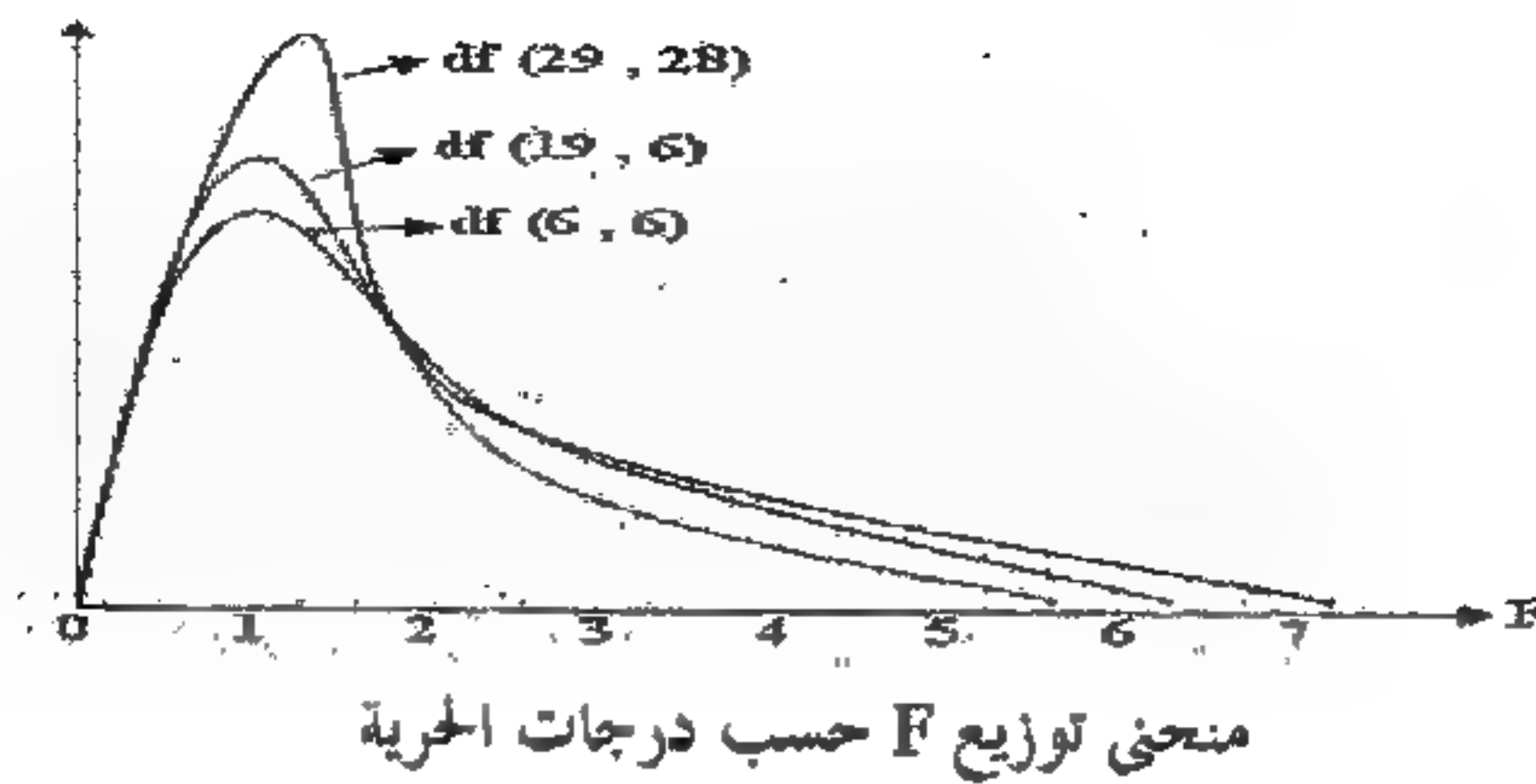
1-9 مقدمة Introduction

تحليل التباين (ANOVA) Analysis Of Variance: هو تعميم لاختبار t ويستخدم لاختبار أكثر من عينتين، ويستخدم لمقارنة عدة متوسطات حسابية في آن واحد $H_0: \mu_1 = \mu_2 = \dots = \mu_k$ تحصيل التباين هو طريقة ذكية لاختبار اختلاف أوساط أكثر من مجموعتين دفعة واحدة من خلال التباين.

يمكن استخدام الرسم البياني Box Plot لتوضيح نتائج المقارنة بين متوسط أكثر من عينتين من العينات المستقلة.

توزيع F عبارة عن مجموعة من المنحنيات التكرارية يتميز كل منها عن الآخر برقمين لدرجات الحرية أحدهما يمثل درجة حرية للبسط والآخر درجة حرية للمقام . وقيمة F هي قيمة توضح نسبة التباين Variance ratio لعينتين والرمز F إشارة إلى العالم Fisher الذي قام بعمل هذا الاختبار والمعروف باختبار F . وقد قام العالم Snedecor بحساب جداول خاصة لتوزيع F وفيها درجات الحرية التي في أعلى الجدول تخص البسط أما درجات الحرية على العمود الجانبي فتخص المقام.

توزيع F : هو توزيع ملئو جهة اليمين بمعلمتين تتمثلان بدرجة حرية (البسط ، المقام) وهما $k - 1$ للبسط ، $k - n$ للمقام حيث n مجموع أحجام العينات، فإذا كان لدينا اختبار لقياس معنوية الفرق بين التقديرين (F) نوجد F_α حيث α مستوى المعنوية المستخدم للفرضية H_0 التي ترفض إذا كان $F < F_\alpha$ وإلا تؤكد بوجود الاختلاف بين المتوسطات، والشكل التالي يبين توزيع F .



وتحليل التباين هو عملية يقصد بها تقسيم مجموع مربعات الانحرافات عن المتوسط الحسابي إلى مكوناته أي إرجاع كل من هذه المكونات إلى مسبباتها. وطريقة تحليل التباين تفيد في مقارنة عدد من المعاملات يزيد عن اثنين. كما تمتاز طريقة تحليل التباين بأنه يمكن فيها استعمال كل البيانات المأخوذة من التجربة في حساب قيمة واحدة للانحراف القياسي يمكن بها مقارنة المجموعات أو المعاملات التجريبية. وكلما زاد عدد المتوسطات كلما زاد احتمال الخطأ وقل احتمال اتخاذ قرار صحيح ففي المثال الخاص بمقارنة متوسطات في ست مناطق فإنه يلزم إجراء الاختبار 15 مرة وبالتالي سوف ينخفض احتمال اتخاذ قرار صحيح في الخمسة عشر اختبار معاً من 0.95 إلى 0.95^{15} أي إلى 0.46 فقط وبالتالي يرتفع احتمال الخطأ في اتخاذ القرار الصحيح من مجرد 0.05 إلى 0.54 والذي يساوي $(1-0.46)$ وهو احتمال كبير جداً للخطأ في اتخاذ القرار.

لذلك كان لابد من التفكير في أسلوب آخر بديل يوفر الوقت والجهد وفي الوقت نفسه لا يقل احتمال اتخاذ القرار الصحيح أو يكبر احتمال الخطأ في اتخاذ القرار، هذا الأسلوب هو الذي يسمى "تحليل التباين" والذي يختبر ما إذا كانت المتوسطات كلها متساوية مرة واحدة دون أخذهم اثنين اثنين ودون أن ينخفض احتمال اتخاذ قرار صحيح أو يزيد احتمال الخطأ عند اتخاذه. وهو الذي يسمى اختصاراً ANOVA وهو اختصار للمصطلح الإنجليزي Analysis of Variance.

ويعتمد هذا الأسلوب من أساليب التحليل الإحصائي على ما يعرف باختبار F والذي يعتمد أساساً على تحليل التباين. أن التباين ما هو إلا متوسط مربعات انحرافات القيم عن

وسطها الحسابي. أي أن التباين يعتمد أساساً على مجموع مربعات ثم القسمة على عدد المشاهدات. ويعتمد أسلوب تحليل التباين على تقسيم مجموع المربعات الكلي إلى أقسام فيمثل كل منها أو يقيس أحد مصادر التغير أو الاختلاف Source of Variation يمثل أحدها - مثلاً - التغير بسبب المعاملات (أو المجتمعات) المختلفة ، ويمثل آخر التغير بسبب الأخطاء ثم تعرف الإحصائية (أو الاختبار) F بأنها خارج قسمة التباين بسبب المعاملات على التباين بسبب الأخطاء وهكذا. أي أنه يتم حساب التباين بسبب المعاملات، والتباين بسبب الأخطاء فيحصل على قيمة F المحسوبة وبمقارنة هذه القيمة بالقيمة الجدولية F نصل إلى قرار إما بقبول الفرضية الصفرية أو عدم قبولها عند مستوى المعنوية المطلوب. ولتحليل التباين تطبيقات كثيرة في مختلف المجالات.

إذا كان هناك متغير مستقل واحد، نسميه تحليل التباين الأحادي One-Way ANOVA

إذا كان هناك متغيرين مستقلين، نسميه تحليل التباين الثنائي Two-Way ANOVA

إذا كان هناك 3 متغيرات مستقلة، نسميه تحليل التباين الثلاثي Three-Way ANOVA ولكن يوجد متغير تابع وحيد في جميع تحليلات التباين أعلاه.

شروط استخدام اختبار تحليل التباين:

1. أن تكون العينات عشوائية مستقلة.
2. أن تكون العينات مسحوبة من مجتمعات لها توزيعات طبيعية.
3. أن تكون تباينات المجتمعات متساوية $\sigma^2_1 = \sigma^2_2 = \sigma^2_3 = \dots = \sigma^2_k = \sigma^2$ ويتم إجراء اختبار تحليل التباين اعتماداً على اختبار F واعداد جدول تحليل التباين.

2-9 تحليل التباين الأحادي One-Way Analysis of Variance

إذا كان لكل فرد من أفراد العينة علامة على متغيرين، الأول يسمى المتغير العامل Factor أو المتغير المستقل Independent Variable وهو متغير من النوع الاسمي Nominal أو الترتيبي Ordinal أما المتغير الآخر فهو المتغير التابع Dependent Variable وهو متغير من النوع الكمي وهو المتغير الذي سيتم فحص مساواة متوسطه لكل فئة من فئات المتغير المستقل.

إن قيم x يتم تصنيفها إلى k من العينات وفقاً لمعيار واحد معين، مثل درجات الطلاب في k من الشعب وكل شعبة تضم n من الطلاب.

يعتبر أبسط التصميمات التجريبية إذ فيه توزع معاملات التجربة Treatments على كل الوحدات التجريبية أو الأفراد عشوائياً . بمعنى أنه إذا كان لدينا في التجربة خمس معاملات مختلفة وكل من المعاملات سوف تنفذ من أربع وحدات فإن استعمال التوزيع العشوائي سوف يحسن لكل مجموعة مكونة من أربع وحدات تجريبية فرصة متساوية لأن تعامل بأي معاملة من معاملات التجربة. أي أن كل وحدة تجريبية توضع في أي معاملة من المعاملات عشوائياً.

مميزات هذا التصميم :

1. يمكن استخدام أي عدد من المعاملات أو المكررات.
2. مفيد في ما إذا كانت الوحدات التجريبية قليلة أو من الصعب التحكم فيها كما في التجارب على الحيوانات.
3. طريقة التحليل بسيطة وسهلة حتى بفرض عدم تساوي المكررات (الوحدات) بين كل المعاملات.
4. تبقى طريقة التحليل الإحصائي بسيطة إذا ما فقدت قيم بعض الوحدات التجريبية أو حتى القيم الكلية لبعض المعاملات.

تحليل التباين الاحادي في حالة تساوي أحجام العينات

إذا كان لدينا عينات عشوائية حجم كل منها n ، مسحوبة من مجتمعات توزيعها طبيعي، ومتوسطاتها $\mu_1, \mu_2, \dots, \mu_k$ ، وتباين σ^2 ، والمطلوب اختبار فرضية:

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k$$

$$H_1: \mu_1 \neq \mu_2 \neq \dots \neq \mu_k$$

لاحتبار مساواة متوسطات المجموعات فإن تحليل التباين يستهدف تجزئة التباين الكلي للمتغير التابع إلى جزئين، ومن ثم تتم المقارنة بين تبايني الجزئين باستخدام اختبار F لذلك عملياً نجزيء مجموع مربعات التباين ودرجات الحرية V إلى تباين الأول معروف المصدر وهو بين المجموعات (Between Groups) ومصدرها الفروقات بين متوسطات المجموعات، فإذا كان

هذا الجزء كبيراً فإن متوسطات المجموعات غير متساوية، والثانية تباين ضمن المجموعات (Within Groups) وهي الجزء غير المعروف مصدره ويمكن أن يسمى الباقي (Residuals) أو الخطأ (Error).

$$SST = SSB + SSW$$

إن صيغة الاحصاء المستخدمة لاختبار الفرضية هي:

$$F = \frac{MSB}{MSW} \dots F_{k-1, n-k}$$

القرار:

يكون القرار رفض الفرضية الصفرية H_0 التي تقول أن متوسطات المجموعات متساوية، إذا كانت نسبة التباين بين المجموعات إلى التباين ضمن المجموعات كبيراً وهذه النسبة تسمى (قيمة F)، فإذا كانت قيمة F المحسوبة أكبر من القيمة الجدولية $F_{\alpha, k-1, k(n-1)}$ نرفض الفرضية الصفرية التي تقول أن متوسطات المجموعات متساوية.

يعتبر هذا التصنيف هو أبسط أنواع تحليل التباين، حيث تصنف المشاهدات إلى عدة مجموعات على أساس متغير واحد أو خاصية واحدة.

والافتراضات الأساسية لهذا التحليل هي ما يلي:

- 1- نفترض أن عدد المجتمعات K وأنها جميعاً مستقلة.
- 2- أنها جميعاً تخضع للتوزيع الطبيعي بمتوسطات تساوي $\mu_1, \mu_2, \dots, \mu_k$
- 3- أن لها جميعاً التباين نفسه σ^2 أي أن التباين لكل المجتمعات ثابت σ^2
- 4- يتم سحب عينة عشوائية من كل من هذه المجتمعات وأن أحجام هذه العينات كلها متساوية وتساوي n ويمكن بكل بساطة افتراض عدم تساوي أحجام العينات ولن يختلف أسلوب التحليل على الإطلاق إلا في أشياء بسيطة جداً.

وتصنف البيانات عادة في هذا التحليل على النحو التالي :

المجتمع أو المعاملة	الملاحظات		
	1	2	n
1	y_{11}	y_{12}	y_{1n}
2	y_{21}	y_{22}	y_{2n}
\vdots	\vdots		
K	y_{k1}	y_{k2}	y_{kn}

حيث يمثل الصف الأول مشاهدات العينة الأولى أي المسحوبة من المجتمع الأول، ويمثل الصف الثاني مشاهدات العينة الثانية المسحوبة من المجتمع الثاني،... وهكذا يمثل الصف الأخير مشاهدات العينة الأخيرة k المسحوبة من المجتمع الأخير رقم K.

كما تكون خطوات الاختبار كما يلي :

1- الفرض الصفري : هو أن متوسطات هذه المجتمعات متساوية، وبالرموز

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_k$$

2- الفرض البديل : هو أن بعض هذه المتوسطات غير متساوية (أو : يوجد متوسطان على الأقل غير متساويين).

3- إحصائية الاختبار : في هذه الحالة يرمز لها بالرمز F وتأخذ الشكل التالي :

$$F = \frac{S_R^2}{S_E^2}$$

والتي لها توزيع F بدرجات حرية للبسط K-1 وللمقام K(n-1) حيث : S_R^2 هو

التباين بسبب الظاهرة أو المتغير أو المعاملات، S_E^2 هو التباين بسبب الخطأ.

ويمكن الحصول على الإحصائية F بتنظيم الحسابات في جدول يسمى " جدول تحليل

التباين " ANOVA TABLE كما يلي :

جدول تحليل التباين (ANOVA)*

مصدر التغير	مجموع المربعات	درجات الحرية	متوسط المربعات	(الإحصائية) المحسوبة
بسبب المعاملات	SSR	K- 1	$S_R^2 = SSR / K - 1$	$F = \frac{S_R^2}{S_E^2}$
بسبب الخطأ	SSE	K (n-1)	$S_E^2 = SSE / K (n-1)$	
الكلية	SST	N*k-1		

وسوف نوضح من المثال كيفية حساب المقادير الثلاثة SSE, SSR, SST

4- حدود منطقتي القبول والرفض: ويتم الحصول عليها من جدول توزيع F بدرجات حرية للبسط K-1 وللمقام k(n-1). (اختبار الطرف الأيمن).

5- المقارنة والقرار: إذا كانت قيمة F المحسوبة من تحليل التباين أقل من قيمة F الجدولية نقبل الفرض العدمي بتساوي المتوسطات والعكس صحيح.

مثال 9-1 : البيانات التالية تمثل أعمار أربع عينات عشوائية من الناخبين سحبت من أربع مدن مستقلة (نفترض أن لها توزيعات طبيعية بمتوسطات $\mu_1, \mu_2, \mu_3, \mu_4$ وتباين مشترك يساوي σ^2):

المدن k	الملاحظات						المجموع
	1	2	3	4	5	6	
الأولى 1	20	21	25	28	30	26	150
الثانية 2	23	22	27	20	26	20	138
الثالثة 3	19	20	21	28	20	18	126
الرابعة 4	24	29	30	28	27	24	162

والمطلوب اختبار الفرض العدمي بأن متوسطات أعمار الناخبين من المدن الأربع متساوية، أي أن المطلوب بالرموز هو :

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$$

وذلك بمستوى معنوية 5%

الحل: تكون خطوات الحل كما يلي:

1- الفرض الصفري:

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$$

2- الفرض البديل: أن بعض هذه المتوسطات غير متساو (أثنان على الأقل غير

متساويين).

3- الإحصائية: وهي في هذه الحالة $F = \frac{S_R^2}{S_E^2}$ وتكون الحسابات التفصيلية لتحليل

التباين كما يلي :

حيث $K = 4$ $n = 6$

أ) متوسطات الصفوف (المدن) :

متوسط الصف الأول : $150/6 = 25$

متوسط الصف الثاني : $138/6 = 23$

متوسط الصف الثالث : $126/6 = 21$

متوسط الصف الرابع : $162/6 = 27$

$$\text{ب) المتوسط الكلي : } \frac{150+138+126+162}{24} = \frac{576}{24} = 24$$

ج) مجموع المربعات الكلي :

$$SST = (20^2 + 21^2 + 25^2 + 28^2 + 30^2 + 26^2 + 23^2 + \dots + 27^2 + 24^2) - 6 \times 4 \times (24)^2$$

$$= 14160 - 13824$$

$$= 14160 - 6 \times 4 \times 24 \times 24$$

$$SST = 336$$

د) مجموع المربعات للصفوف (المدن) :

$$SSR = 6(25^2 + 23^2 + 21^2 + 27^2) - 6 \times 4 \times (24)^2$$

$$= 13944 - 13824$$

$$SSR = 120$$

هـ) مجموع مربعات الخطأ :

$$SSE = SST - SSR$$

$$= 336 - 120$$

$$SSE = 216$$

و) ثم نكون جدول تحليل التباين كما يلي : $k=4, n=6$

مصدر التغير	مجموع المربعات	درجات الحرية	متوسط المربعات	الإحصائية F
بسبب المعاملات	SSR= 120	$k-1$ 4-1 3	$S^2R = 120/3 = 40$	$F = \frac{40}{10.8} = 3.7$
بسبب الخطأ	SSE= 216	$K(n-1)$ 4(6-1) 20	$S^2E = 216/20 = 10.8$	
الكلية	SST= 336	$Nk-1$ 6*4-1 23		

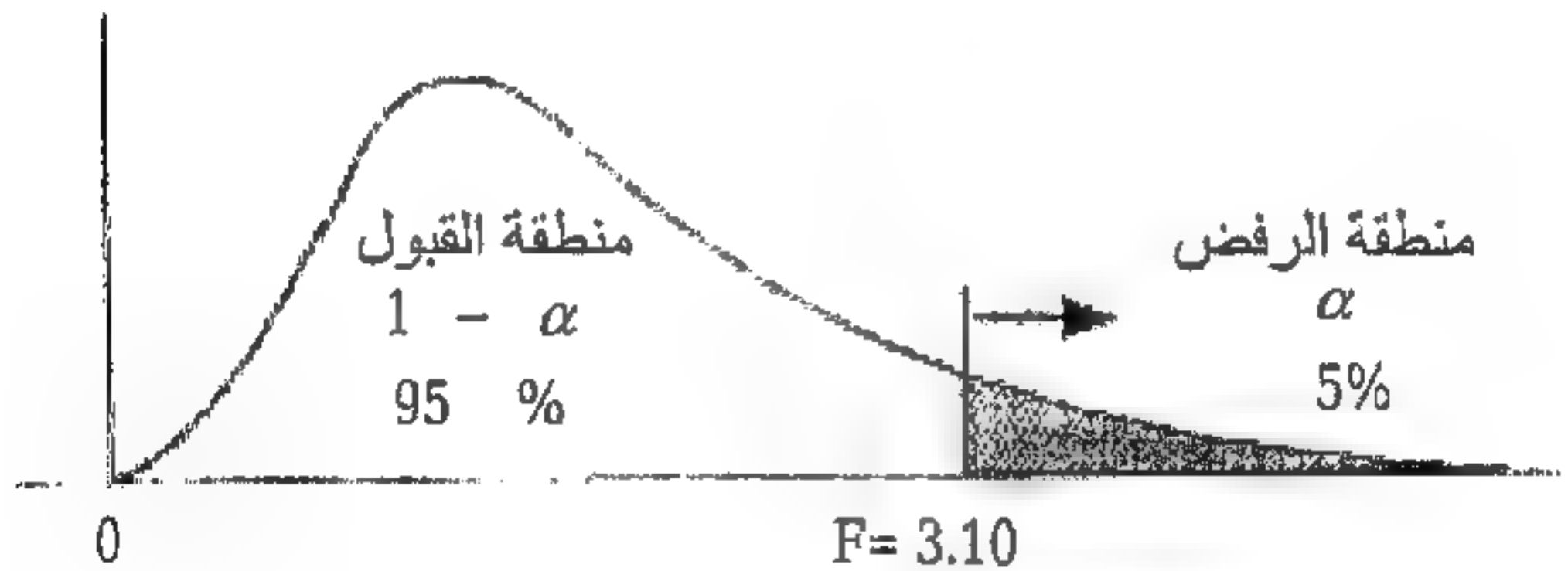
أي أن قيمة الإحصائية (أو F المحسوبة) هي 3.7

$$F = \frac{40}{10.8} = 3.7$$

4- حدود منطقتي القبول والرفض : من جدول توزيع F وعند مستوى معنوية 5%

وبدرجات حرية 3 للبسط، 20 للمقام نجد أن F الجدولية تساوي 3.10

ويمكن توضيح ذلك بالرسم كما يلي :



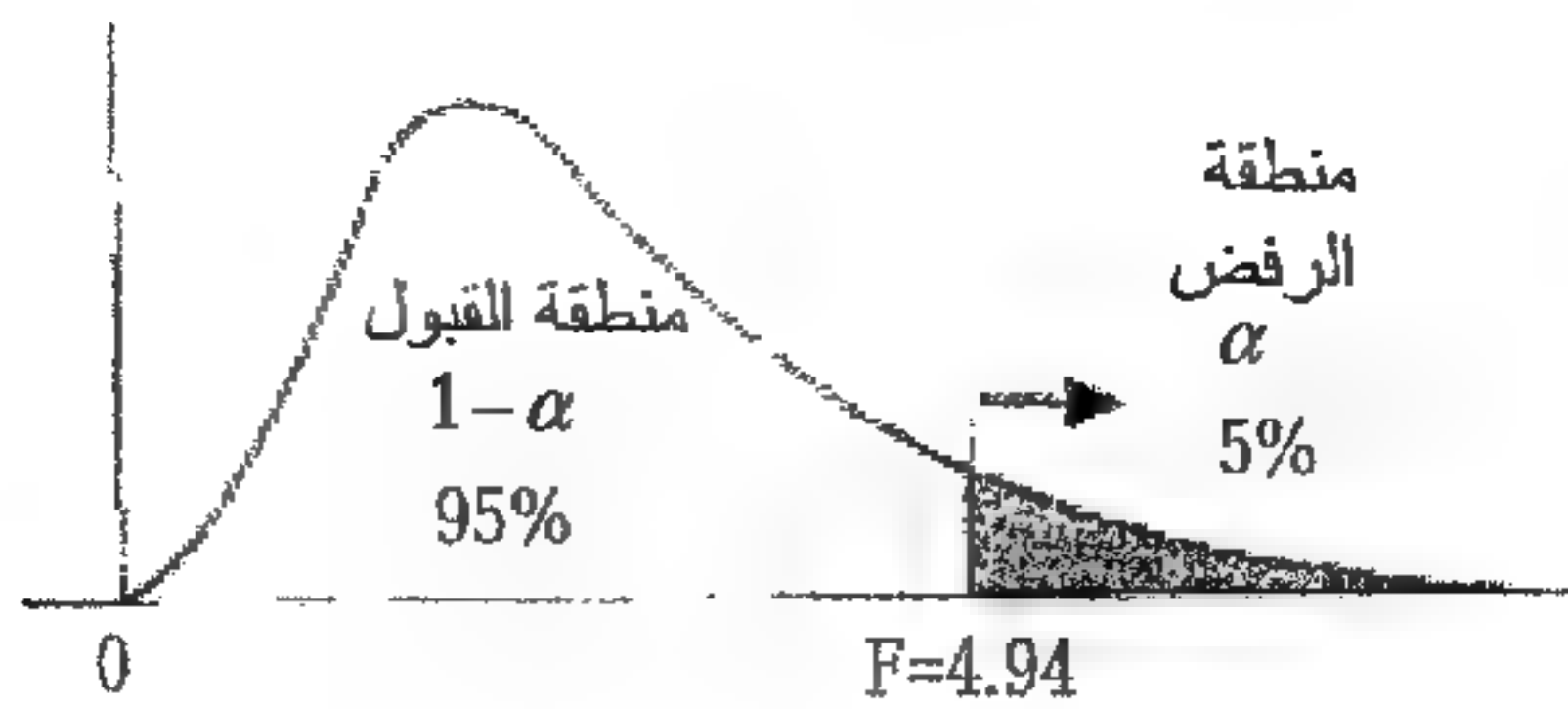
5- المقارنة والقرار: وحيث أن قيمة الإحصائية المسحوبة والتي تساوي 3.7 أكبر من

القيمة الجدولية فإنها تقع في منطقة الرفض وبالتالي فإن القرار هو رفض الفرض العدمي بتساوي

متوسطات أعمار الناحيين في المدن الأربع وذلك بمستوى معنوية 5%

* أما إذا استخدمنا مستوى معنوية 1% فإن قيمة F من الجدول تصبح 4.94 أي تصبح

حدود منطقتي القبول والرفض كما يلي :



في هذه الحالة فإن قيمة الإحصائية والتي تساوي 3.7 تقع في منطقة القبول، وبالتالي فإن القرار يكون قبول الفرض العدمي ينساوي متوسطات أعمار الناجحين في المدن الأربع وذلك بمستوى معنوية 1%.

مثال 9-2: في دراسة لأثر طريقة التدريس في التحصيل في مادة الاحصاء لدى طلبة الجامعة. باستخدام ثلاثة طرق. كما هو موضح في الجدول أدناه:

الطريقة الأولى	الطريقة الثانية	الطريقة الثالثة
2	5	7
3	6	5
4	5	6
5	7	8
6	7	9

هل هناك فرق ذو دلالة احصائية في التحصيل يعزى لطريقة التدريس.

الحل:

متغيرات الدراسة:

المتغير المستقل: طريقة التدريس وتقسم إلى: الطريقة 1

الطريقة 2

الطريقة 3

المتغير التابع: التحصيل في مادة الاحصاء، ويقاس بالعلامة على الاختبار التحصيلي.

فرضيات الدراسة:

$H_0: \mu_1 = \mu_2 = \mu_3$ الفرضية الصفرية

$H_1: \mu_1 \neq \mu_2 \neq \mu_3$ الفرضية البديلة

الاقتراح: تجزئة الفرضية الصفرية الأم الى مجموعة من الفرضيات البسيطة واستخدام

اختبار t في كل مرة، ويكون عدد الفرضيات الجزئية $n(n-1)/2$

الاقتراح مرفوض للأسباب التالية:

1. لا يمكن من خلال اختبار t دراسة أثر التفاعل.
2. غير عملي في ظل زيادة عدد المتوسطات الحسابية.
3. تقل قيمة قوة الاختبار مع تثبيت قيمة α .
4. تزداد قيمة α (زيادة احتمال الوقوع في الخطأ) وهو شيء غير مرغوب فيه.

	الطريقة 1		الطريقة 2		الطريقة 3			
	G1	G1 ²	G2	G2 ²	G3	G3 ²	Total	Total ²
	2.00	4.00	5.00	25.00	7.00	49.00	2	4
	3.00	9.00	6.00	36.00	5.00	25.00	3	9
	4.00	16.00	5.00	25.00	6.00	36.00	4	16
	5.00	25.00	7.00	49.00	8.00	64.00	5	25
	6.00	36.00	7.00	49.00	9.00	81.00	6	36
							5	25
المجموع	20.00		30.00		35.00		6	36
المتوسط	4.00		6.00		7.00		5	25
Σx^2	90.00		184.00		255.00		7	49
العدد	5		5		5		7	49
							7	49
							5	25
							6	36
							8	64
							9	81
المجموع							85	
المتوسط							5.67	
Σx^2							529	
عدد							15	

لاختبار الفرضية الصفرية $H_0: \mu_1 = \mu_2 = \mu_3$

1- نحسب التباين بين المجموعات (Between Groups) وهو (MSB)

2- نحسب التباين ضمن المجموعات (Within Groups) وهي (MSW)

3- نحسب صيغة الاحصاء المستخدمة لاختبار الفرضية هي:

$$F = \frac{MSB}{MSW} \dots\dots F_{k-1, n-k}$$

درجة الحرية للبسط = (عدد المجموعات - 1)

درجة الحرية للمقام = (حجم العينة - عدد المجموعات)

4- القرار: إذا كانت قيمة F المحسوبة أكبر من القيمة الجدولية $F_{\alpha, k-1, k(n-1)}$

نرفض الفرضية الصفرية التي تقول ان متوسطات المجموعات متساوية.

أي تباين في النهاية يصمم على شكل جدول كالتالي:

Source of Variance	Sum of Squares (SS)	Degrees of freedom (DF)	Mean. Squares (MS)	F
مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	ف
Between Groups بين المجموعات	SSB مجموع المربعات بين	C - 1 عدد الاعمدة - 1	MSB=SSB/C-1 متوسط المربعات بين	MSB/MS W
Within Groups ضمن المجموعات	SSW مجموع المربعات ضمن	N - C حجم العينة - عدد الاعمدة	MSW=SSW/N-C متوسط المربعات ضمن	
Total الكل	SST مجموع المربعات الكل	N - 1 حجم العينة - 1		

التباين داخل المجموعات هو تباين عشوائي (ناتج عن اختلاف الأفراد).

التباين بين المجموعات هو تباين عشوائي (ناتج عن اختلاف الأفراد)، وتباين ناتج عن

اختلاف طريقة التدريس.

عندما يكون التباين لطريقة التدريس = 0، فإن التباين الداخلي = التباين الخارجي، وهذا

يحدث في حالة أن الفرضية الصفرية صحيحة.

حجم العينة N ، عدد المجموعات في العينة C

$$MSW = SSW / N - C$$

$$MSB = SSB / C - 1$$

$$SST = SSB + SSW$$

احسب الكميات:

طريقة 1:

$$A = \sum x^2$$

$$B = (\sum x)^2 / N$$

$$C = (\sum x)^2 / n_1 + (\sum x)^2 / n_2 + \dots$$

$$SST = A - B$$

$$SSB = C - B$$

$$SSW = A - C$$

طريقة 2:

$$SST = \sum (X - X')^2$$

بمجموع مربعات الانحراف عن الوسط الكلي

$$SSB = \sum n_i (X'_i - X')^2$$

$$SSW = \sum (X - X'_i)^2$$

لكل المجموعات

طريقة 3: اذا علمت المتوسطات الحسابية والانحرافات المعيارية.

$$SSB = n_1 (X'_1 - X')^2 + n_2 (X'_2 - X')^2 + \dots$$

$$SSW = (n_1 - 1)S_1^2 + (n_2 - 1)S_2^2 + \dots$$

حل المثال 2-9

$$A = \sum x^2$$

$$= 529$$

$$B = (\sum x)^2 / N$$

$$= (85)^2 / 15 = 481.67$$

$$C = (\sum x)^2 / n_1 + (\sum x)^2 / n_2 + \dots$$

$$= (20)^2 / 5 + (30)^2 / 5 + (35)^2 / 5 = 505$$

$$SST = A - B$$

$$= 529 - 481.67 = 47.33$$

$$SSB = C - B$$

$$= 505 - 481.67 = 23.33$$

$$SSW = A - C$$

$$= 529 - 505 = 24$$

Source of Variance	Sum of Squares (SS)	Degrees of freedom (DF)	Mean. Sequares (MS)	F
مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	ف
Between Groups بين المجموعات	SSB 23.33	C - 1 3-1 = 2	MSB=SSB/C-1 23.33/2= 11.665	MSB/MS W 11.665/2 = 5.8325
Within Groups ضمن المجموعات	SSW 24	N - C 15-3 = 12	MSW=SSW/N-C 24/12 = 2	
Total الكل	SST 47.33	N - 1 15-1 = 14		

القرار:

القيمة المحسوبة لـ $F = 5.8325$

القيمة الحرجة (الجدولية) $F_{2,12,0.05} = 3.89$

بما أن القيمة المحسوبة $>$ أكبر من القيمة الحرجة نرفض الفرضية الصفرية، ونستنتج أنه يوجد اختلاف في التحصيل يعزى إلى طريقة التدريس.

ثم سنخرج الجدول للعام للنسبة المئوية

مصدر التباين	مجموع المربعات	درجة الحرية	التباين	النسبة المئوية المئوية
بين المجموعات	41.17	3	147.06	
			+	14.29
ضمن المجموعات	164.63	16	10.29	
تجميع	605.8	19		

ف=14.29، دج=16/3، دالة عند 0.01 (الرجوع إلى جدول النسبة المئوية للقراءات)

كيفية قراءة جدول النسبة الحادية حسب النسبة الحادية التي حصلنا عليها ونسب درجات الحرية التي لدينا.

TABLE XI

TABLE DE F REVERSÉE

Les deux valeurs de F correspondent respectivement aux tests de P = 5% et P = 1%

v	F									
	1	2	3	4	5	6	8	12	20	∞
1	161.45	19.16	10.13	7.71	6.58	5.95	5.19	4.10	3.01	2.58
2	18.51	9.00	5.99	4.74	4.05	3.57	3.01	2.29	1.78	1.58
3	10.13	5.55	3.78	3.12	2.71	2.41	2.04	1.57	1.19	1.08
4	7.71	4.10	2.71	2.29	2.04	1.78	1.58	1.19	0.90	0.82
5	6.58	3.57	2.41	2.04	1.78	1.58	1.19	0.90	0.71	0.64
6	5.95	3.01	2.04	1.78	1.58	1.19	0.90	0.71	0.54	0.48
8	5.19	2.29	1.58	1.19	0.90	0.71	0.54	0.48	0.35	0.30
12	4.10	1.78	0.90	0.71	0.54	0.48	0.35	0.30	0.21	0.18
20	3.01	1.19	0.54	0.48	0.35	0.30	0.21	0.18	0.13	0.11
∞	2.58	1.08	0.48	0.42	0.30	0.25	0.18	0.15	0.10	0.08
1	161.45	19.16	10.13	7.71	6.58	5.95	5.19	4.10	3.01	2.58
2	18.51	9.00	5.99	4.74	4.05	3.57	3.01	2.29	1.78	1.58
3	10.13	5.55	3.78	3.12	2.71	2.41	2.04	1.57	1.19	1.08
4	7.71	4.10	2.71	2.29	2.04	1.78	1.58	1.19	0.90	0.82
5	6.58	3.57	2.41	2.04	1.78	1.58	1.19	0.90	0.71	0.64
6	5.95	3.01	2.04	1.78	1.58	1.19	0.90	0.71	0.54	0.48
8	5.19	2.29	1.58	1.19	0.90	0.71	0.54	0.48	0.35	0.30
12	4.10	1.78	0.90	0.71	0.54	0.48	0.35	0.30	0.21	0.18
20	3.01	1.19	0.54	0.48	0.35	0.30	0.21	0.18	0.13	0.11
∞	2.58	1.08	0.48	0.42	0.30	0.25	0.18	0.15	0.10	0.08
1	161.45	19.16	10.13	7.71	6.58	5.95	5.19	4.10	3.01	2.58
2	18.51	9.00	5.99	4.74	4.05	3.57	3.01	2.29	1.78	1.58
3	10.13	5.55	3.78	3.12	2.71	2.41	2.04	1.57	1.19	1.08
4	7.71	4.10	2.71	2.29	2.04	1.78	1.58	1.19	0.90	0.82
5	6.58	3.57	2.41	2.04	1.78	1.58	1.19	0.90	0.71	0.64
6	5.95	3.01	2.04	1.78	1.58	1.19	0.90	0.71	0.54	0.48
8	5.19	2.29	1.58	1.19	0.90	0.71	0.54	0.48	0.35	0.30
12	4.10	1.78	0.90	0.71	0.54	0.48	0.35	0.30	0.21	0.18
20	3.01	1.19	0.54	0.48	0.35	0.30	0.21	0.18	0.13	0.11
∞	2.58	1.08	0.48	0.42	0.30	0.25	0.18	0.15	0.10	0.08
1	161.45	19.16	10.13	7.71	6.58	5.95	5.19	4.10	3.01	2.58
2	18.51	9.00	5.99	4.74	4.05	3.57	3.01	2.29	1.78	1.58
3	10.13	5.55	3.78	3.12	2.71	2.41	2.04	1.57	1.19	1.08
4	7.71	4.10	2.71	2.29	2.04	1.78	1.58	1.19	0.90	0.82
5	6.58	3.57	2.41	2.04	1.78	1.58	1.19	0.90	0.71	0.64
6	5.95	3.01	2.04	1.78	1.58	1.19	0.90	0.71	0.54	0.48
8	5.19	2.29	1.58	1.19	0.90	0.71	0.54	0.48	0.35	0.30
12	4.10	1.78	0.90	0.71	0.54	0.48	0.35	0.30	0.21	0.18
20	3.01	1.19	0.54	0.48	0.35	0.30	0.21	0.18	0.13	0.11
∞	2.58	1.08	0.48	0.42	0.30	0.25	0.18	0.15	0.10	0.08

خطية التباين

خطية التباين

حدود دلالة عند 0.05
أي لسطر الأول

حدود دلالة عند 0.01
أي لسطر الثاني

3-9 تحليل التباين الثنائي Tow-Way Analysis of Variance

تحليل التباين الثنائي Tow-Way Analysis of Variance:

يستخدم لدراسة أثر متغيرين عاملين يقسم كل منهما أفراد العينة إلى مستويين أو أكثر على متغير كمي (المتغير التابع).

يستخدم لمقارنة عدة متوسطات حسابية بهدف دراسة أثر متغيرين مستقلين على متغير تابع واحد.

أي يستخدم تحليل التباين الثنائي لدراسة الأثر لعاملين هما العامل a والعامل b على متغير تابع واحد، مثل دراسة أثر طريقة التدريس (العامل a)، والجنس (العامل b)، على التحصيل (المتغير التابع).

ويقصد بأثر التفاعل $a \times b$ على المتغير التابع هو اختلاف تأثير a على المتغير التابع من مستوى إلى آخر من مستويات b ، أي يختلف تأثير طريقة التدريس على التحصيل لدى الذكور عنه لدى الإناث، وفحص التفاعل مهم جداً حيث يعني وجوده التخصيص في النتائج، ويعني غيابه تعميم النتيجة.

إن تحليل التباين الأحادي يستخدم لدراسة أثر عامل واحد (المتغير العامل) على متغير ما. ولكن ماذا لو أردنا دراسة أثر عاملين أو أكثر على متغير ما؟ في هذه الحالة يمكننا استخدام تحليل التباين الثنائي أو الثلاثي.

كثيراً ما يواجه الباحثون مشاكل تتطلب دراسة نوعين من العوامل أو المعالجات وفي هذه الحالة يتم إجراء تحليل التباين بوجود معيارين للتصنيف فمثلاً يكون المعيار الأول (طلاب في مدرسة) والمعيار الثاني (طرق تدريس مختلفة) فإذا أردنا معرفة ما إذا كان الاختلاف بين طلاب المدرسة عائد على اختلاف طريقة التدريس المتبعة معهم، وقد نرغب في معرفة التأثير المشترك لكلا العاملين (أي تأثير التفاعل). وفي الدراسات التطبيقية التجريبية يمكن مثلاً دراسة تأثير التربة ونوعية السماد المستخدم في إنتاج القمح، أو دراسة تأثير جودة مواد البناء ونوعية المهندسين لعمل البيوت السكنية، أو دراسة تأثير مناطق بيع البضائع ومصاريف الدعاية على كمية المبيعات.

فتحليل التباين الثنائي **Two Way ANOVA** يمكن استخدامه لدراسة اثر متغيرين عاملين يقسم كل منهما مفردات العينة الى مستويين (مجموعتين) او أكثر على متغير كمي ما (المتغير التابع). ويستخدم في هذه الحالة جداول مزدوجة مكونة من صفوف وأعمدة وتضم كل خلية من خلاياها مشاهدتين أو أكثر (مفردات تنتمي إلى صف معين وعمود بنفس الوقت).

• إن تحليل التباين باتجاهين يستخدم لمعرفة إن كانت هناك فروق معنوية بين المتوسطات (متوسطات معالجات الصفوف، بين متوسطات معالجات الأعمدة، وكذلك بين متوسطات المشاهدات الموجودة في كل خلية حيث تستخدم متوسطات الخلايا لاختبار التفاعل).

من خلال تحليل التباين الثنائي يمكن اختبار ثلاث فرضيات كما يلي:

1. الأثر الرئيسي للمتغير العامل الأول على المتغير التابع. $H_0: \mu a_1 = \mu a_2 = \mu a_3$
 2. الأثر الرئيسي للمتغير العامل الثاني على المتغير التابع. $H_0: \mu b_1 = \mu b_2 = \mu b_3$
 3. أثر التفاعل بين المتغيرين العاملين على المتغير التابع. $a*b$
- $$H_0: \sum (\alpha B)^2 = 0$$

شروط تحقيق التباين الثنائي:

- يجب أن يكون توزيع المتغير التابع طبيعياً لكل مجتمع من المجتمعات في تصميم التجربة، أي أن كل مجتمع ممثل بكل خلية من خلايا تصميم التجربة، فإذا كان على سبيل المثال ثلاث مستويات لكل متغير عامل فيكون هناك 9 خلايا. وإن لم يتحقق هذا الشرط فإنه يمكن الاستغناء عنه بزيادة حجم العينة بحيث تزيد على 15 مفردة لكل مجموعة (خلية)، وفي هذه الحالة قد تكون نتيجة تحليل التباين دقيقة إلى حد ما حتى لو كان توزيع المتغير التابع ليس طبيعياً.
- يجب أن يكون تباين المتغير التابع متساوياً لكل مجتمع من المجتمعات المعرفة في كل خلية من خلايا تصميم التجربة، وإذا لم يتحقق هذا الشرط فإن نتيجة تحليل التباين لن تكون دقيقة. أما المقارنات البعدية الخاصة بالأثر الرئيسي فمن الممكن استخدام بعض الطرائق التي لا تشترط تساوي التباين.
- يجب أن تكون العينات مختارة بطريقة عشوائية من كل مجتمع من المجتمعات. ويجب أن تكون قيم المتغير التابع مستقلة عن بعضها بعضاً لكل مفردة من مفردات العينات.

هناك نماذج متعددة لتحليل التباين الثنائي ترتبط بالتصميم التجريبي الذي يعتمد عليه الباحث منها:

1. تحليل التباين بآثار ثابتة Fixed Effects
2. تحليل التباين بآثار عشوائية Random Effects
3. تحليل التباين بآثار خليطة Mixed Effects

تحليل التباين بآثار ثابتة Fixed Effects

مثال:

- 1- الجنس: وله مستويان ذكر وانثى.
 - 2- طرق تعليم القراءة: ولها ثلاثة مستويات.
- المتغير التابع : القدرة على الاستيعاب.

* تحليل التباين الثنائي Two Way Analysis of Variance

تحليل التباين الثنائي Two Way ANOVA يمكن استخدامه لدراسة اثر متغيرين عامدين يقسم كل منهما مفردات العينة الى مستويين (مجموعتين) او اكثر على متغير كمي ما (المتغير التابع).

مثال:

باستخدام ملف Employee data الموجود مع برمجية SPSS احتر الفرضية التالية:
 " لا يؤثر الجنس ونوع العمل في تحديد الراتب للموظفين بمستوى معنوية 0.05 "

الحل:

المتغيرات المستقلة: إن هناك عاملان يؤثران على تحديد الراتب هما:

- 1- الجنس.
- 2- نوع العمل.

المتغير التابع: الراتب.

لذلك يمكن تقسيم هذه الفرضية الى ثلاث فرضيات جزئية وهي:

الفرضية الاولى " لا تأثير للجنس في تحديد الراتب "

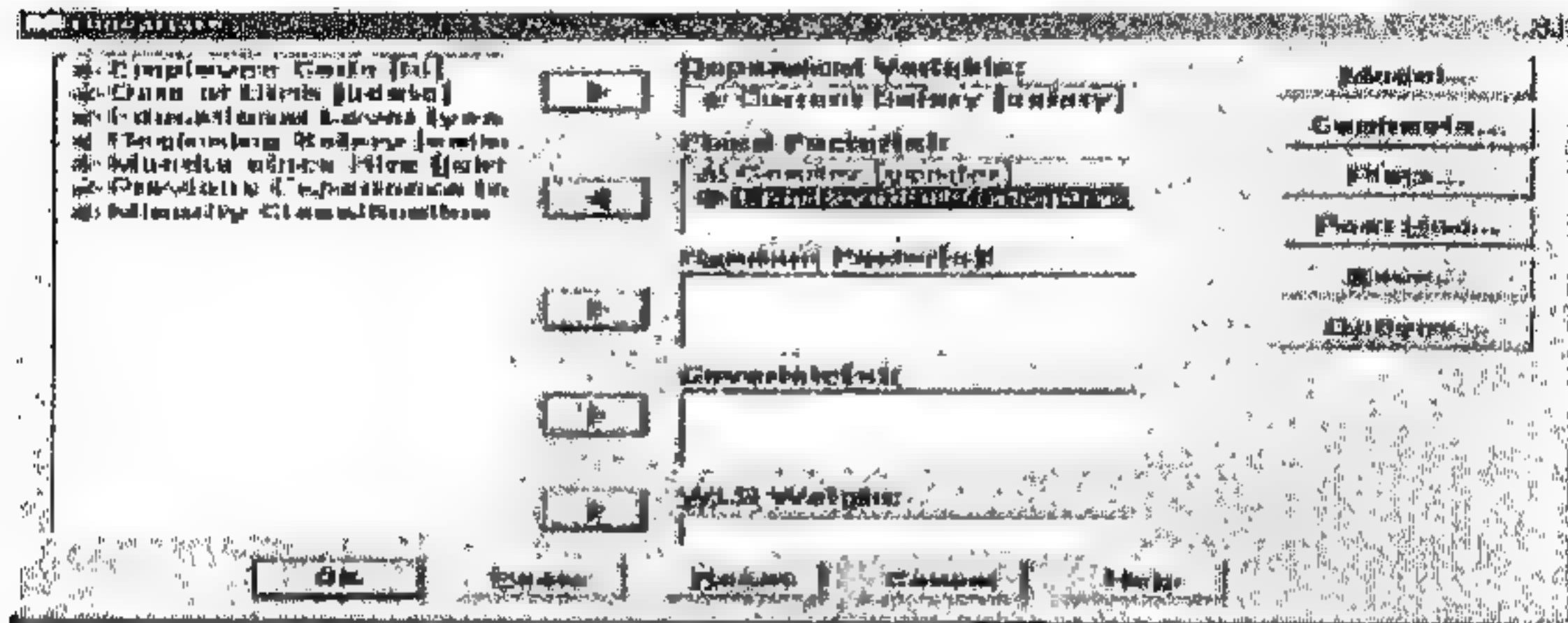
الفرضية الثانية " لا تأثير لنوع العمل على تحديد الراتب "

الفرضية الثالثة " لا يوجد تفاعل بين متغير الجنس ومتغير نوع العمل "

ولفحص الفرضيات نستخدم تحليل التباين الثنائي كما يلي:

1. من القائمة Analyze اختر General Linear Model ومن القائمة الفرعية اختر

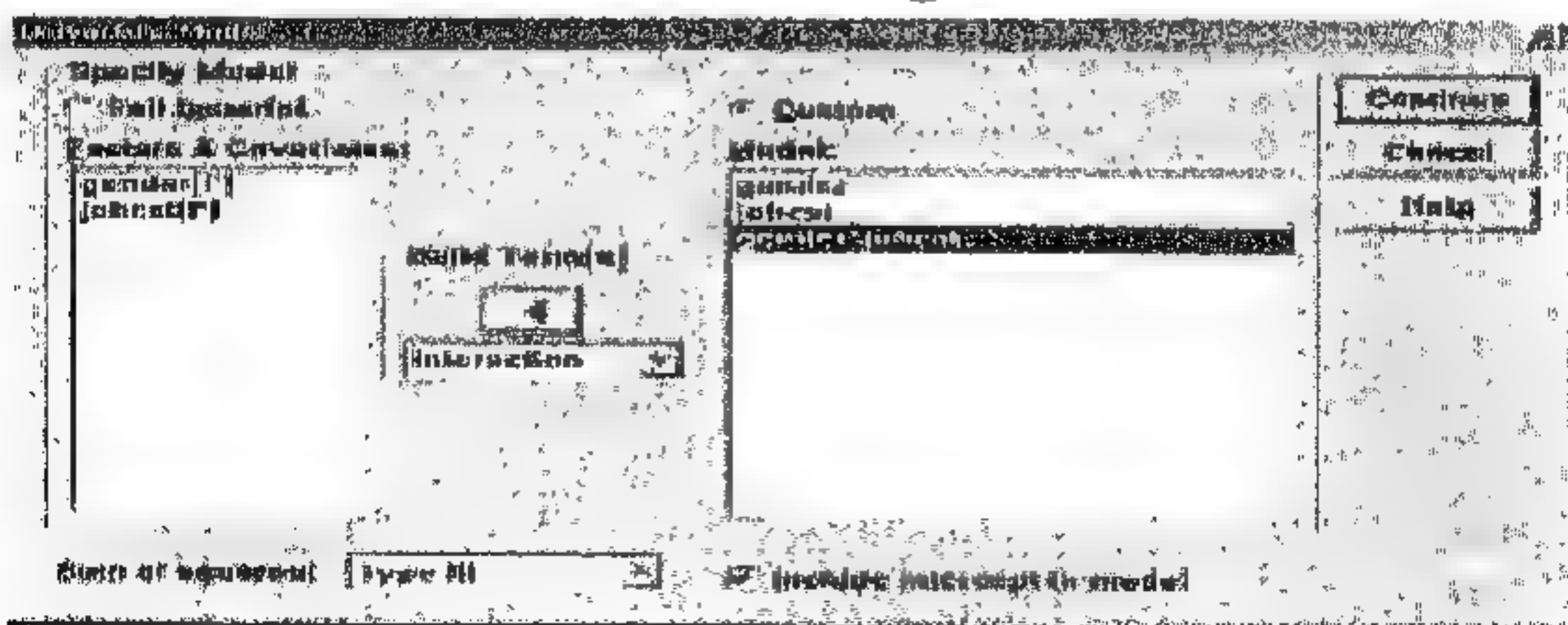
Univariate يظهر مربع الحوار التالي:



2. انقل المتغير Salary الى المستطيل أسفل Dependent Variable والمتغيران

Gender و Jobcat الى المستطيل أسفل Fixed Factor(s) .

3. اضغط على Model يظهر مربع الحوار التالي:



4. اختر Custom ثم Main effects من القائمة أسفل Build Term(s) وانقل

المتغيرين Gender و Jobcat الى المستطيل أسفل Model ، ثم اختر Interaction من القائمة

Build Term(s) وانقل المتغيرين Gender و Jobcat معا الى المستطيل أسفل Model ،

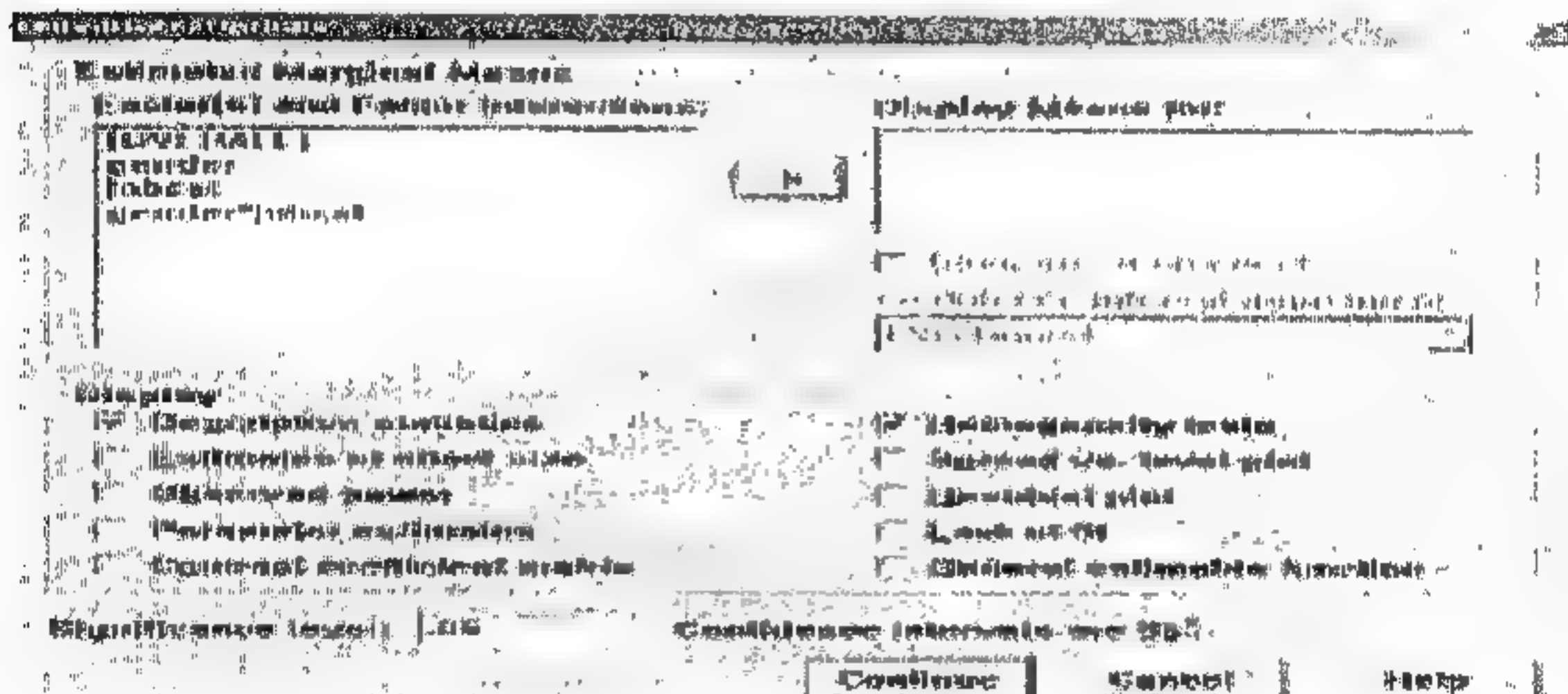
اضغط Continue سنعود الى المربع الاصلي.

5. اضغط Options سيظهر مربع الحوار التالي اختر منه Descriptive statistics

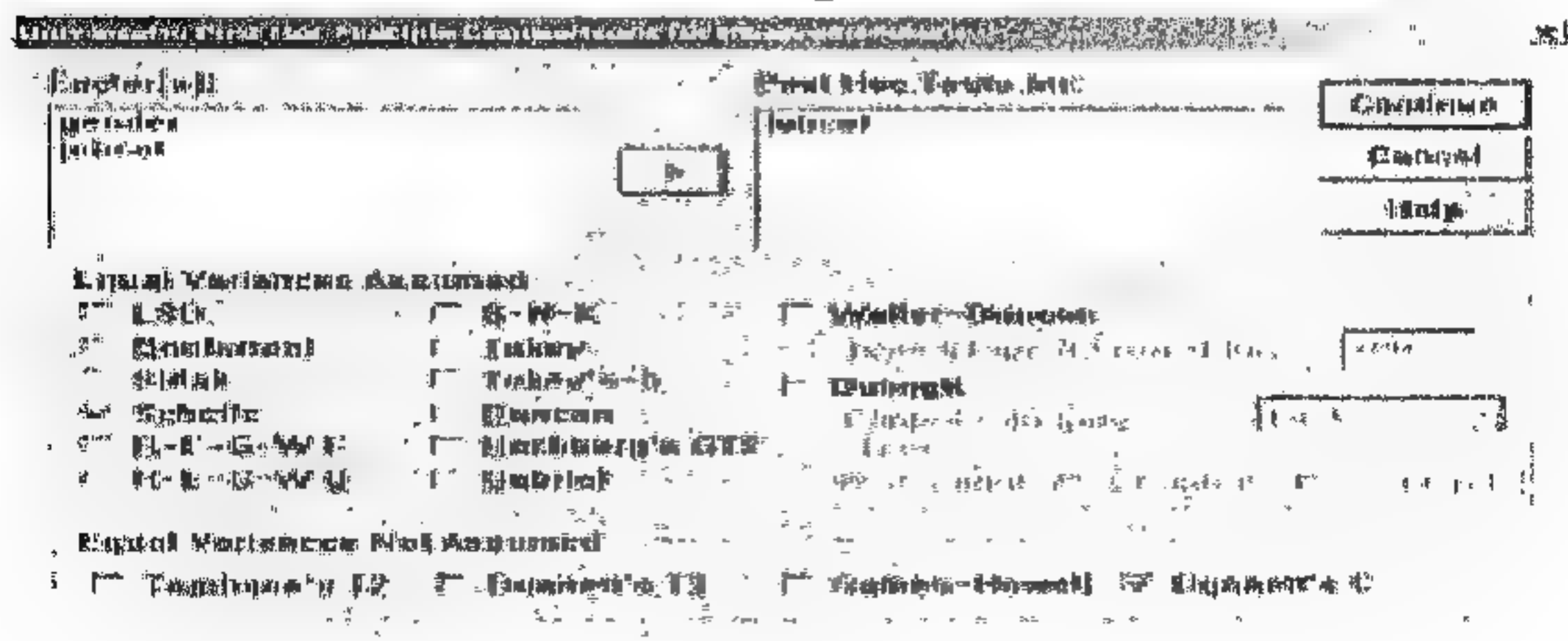
والخيار Homogeneity tests ثم اضغط Continue لنعود لمربع الحوار الأصلي.

Descriptive Statistics

Gender	Employment Category	Mean	Std. Deviation	N
Female	Clerical	\$25,003.69	\$5,812.838	206
	Manager	\$47,213.50	\$8,501.253	10
	Total	\$26,031.92	\$7,558.021	216
Male	Clerical	\$31,558.15	\$7,997.978	157
	Custodial	\$30,938.89	\$2,114.616	27
	Manager	\$66,243.24	\$18,051.570	74
	Total	\$41,441.78	\$19,499.214	258
Total	Clerical	\$27,838.54	\$7,567.995	363
	Custodial	\$30,938.89	\$2,114.616	27
	Manager	\$63,977.80	\$18,244.776	84
	Total	\$34,419.57	\$17,075.661	474



6. اضغط على Post Hoc ليظهر مربع الحوار التالي:



7. اختبر اختبار شفیه Scheffe للمقارنات البعدية من قائمة الاختبارات البعدية التي تشترط تماثل تباينات الفئات Equal Variance Assumed .
8. اختبر اختبار دونت س Dunnett,s C من قائمة الاختبارات البعدية التي لا تشترط تماثل تباينات الفئات Equal Variance Not Assumed
9. انقل المتغير Jobcat فقط الى المستطيل اسفل Post Hoc Tests For لأنه يتكون من ثلاث مستويات أما متغير Gender فلا ننقله لأنه يتكون من مستويين فقط.
10. اضغط Continue سنعود لمربع الحوار الأصلي.

Univariate Analysis of Variance

Between-Subjects Factors			
		Value Label	N
Gender	f	Female	216
	m	Male	258
Employment Category	1	Clerical	363
	2	Custodial	27
	3	Manager	84

الجدول التالي يبين توزيع العينة حسب مستويات كل من المتغيرات العائلية.

جدول اختبار تجانس التباين test of Homogeneity of Variances ، ويبين أن قيمة Sig. = 0.0 يعني أن تباين المجموعات غير متساو لأنها اكبر من 0.05.

الجدول التالي يبين تحليل التباين الثنائي حسب فئات المتغير Gender ، ويظهر أن Sig. = 0.0 وهي اقل من 0.05 أي أن الجنس يؤثر في تحديد الراتب. كذلك الجدول التالي يبين تحليل التباين الثنائي حسب فئات المتغير Jobcat ، ويظهر أن Sig=0.0 وهي اقل من 0.05 أي أن نوع العمل يؤثر في تحديد الراتب. كما يظهر أن هناك تفاعل بين الجنس ونوع الوظيفة لان قيمة Sig=0.0 وهي اقل من 0.05

Levene's Test of Equality of Error Variances ^a

Dependent Variable: Current Salary

F	df1	df2	Sig.
33.383	4	469	.000

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+GENDER+JOB CAT+GENDER
*JOB CAT

Multiple Comparisons

Dependent Variable: Current Salary

	(I) Employment Category	(J) Employment Category	Mean Difference (I-J)	Std. Error	Sig.
Scheffe	Clerical	Custodial	-\$3,100.35	\$1,875.539	.256
	Clerical	Manager	-\$36,139.26*	\$1,138.387	.000
	Custodial	Clerical	\$3,100.35	\$1,875.539	.256
	Custodial	Manager	-\$33,038.91*	\$2,080.027	.000
	Manager	Clerical	\$36,139.26*	\$1,138.387	.000
	Manager	Custodial	\$33,038.91*	\$2,080.027	.000
Dunnnett C	Clerical	Custodial	-\$3,100.35*	\$568.679	
	Clerical	Manager	-\$36,139.26*	\$2,029.912	
	Custodial	Clerical	\$3,100.35*	\$568.679	
	Custodial	Manager	-\$33,038.91*	\$2,031.840	
	Manager	Clerical	\$36,139.26*	\$2,029.912	
	Manager	Custodial	\$33,038.91*	\$2,031.840	

Based on observed means.

*. The mean difference is significant at the .05 level.

الجدول التالي يبين أن متوسطات Jobcat مختلفة ويبين أن متوسطات الكتاب المدرس وكذلك متوسطات الحراس والمدرّس مختلفة بينما لا يوجد خلاف بين متوسطات رواتب الحراس والكتاب له دلالة إحصائية تذكر

Tests of Between-Subjects Effects

Dependent Variable: Current Salary

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	9.646E+10 ^a	4	2.411E+10	272.780	.000
Intercept	1.773E+11	1	1.773E+11	2005.313	.000
GENDER	5247440732	1	5247440732	59.359	.000
JOB CAT	3.232E+10	2	1.616E+10	182.782	.000
GENDER * JOB CAT	1247682867	1	1247682867	14.114	.000
Error	4.146E+10	469	88401147.44		
Total	6.995E+11	474			
Corrected Total	1.379E+11	473			

a. R Squared = .699 (Adjusted R Squared = .697)

* تحليل التباين الثلاثي Three Way ANOVA

إذا كان لدينا ثلاث متغيرات عاملية و اردنا فحص اثر هذه العوامل على متغير تابع
نستخدم تحليل التباين الثلاثي و تتبع نفس خطوات تحليل التباين الثنائي و لناخذ المثال التالي:

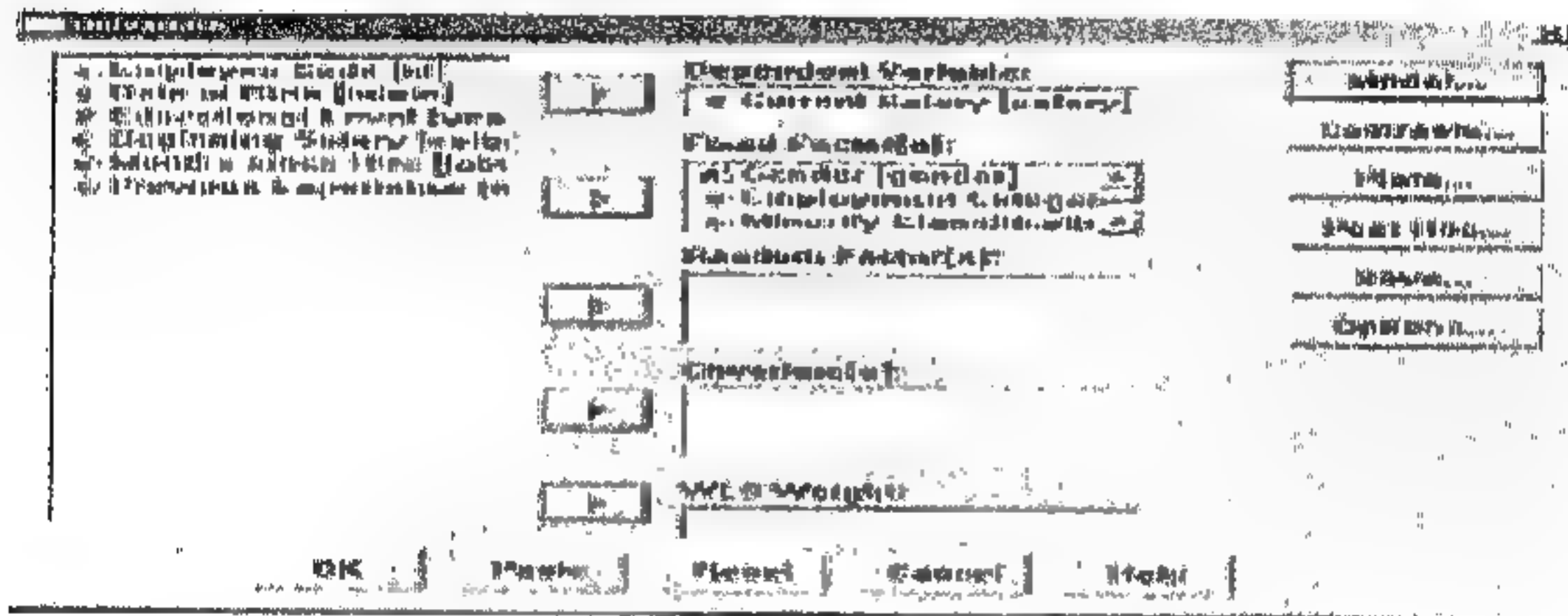
مثال : استخدم ملف Employee data لفحص الفرضية التالية :

" لا يوجد فرق في متوسطات رواتب الموظفين تحت تأثير الجنس و نوع الوظيفة و الاقلية
بمستوى دلالة 0.05 "

ولفحص هذه الفرضية تتبع الخطوات التالية:

1. من القائمة Analyze اختر General Linear Model ومن القائمة الفرعية اختر

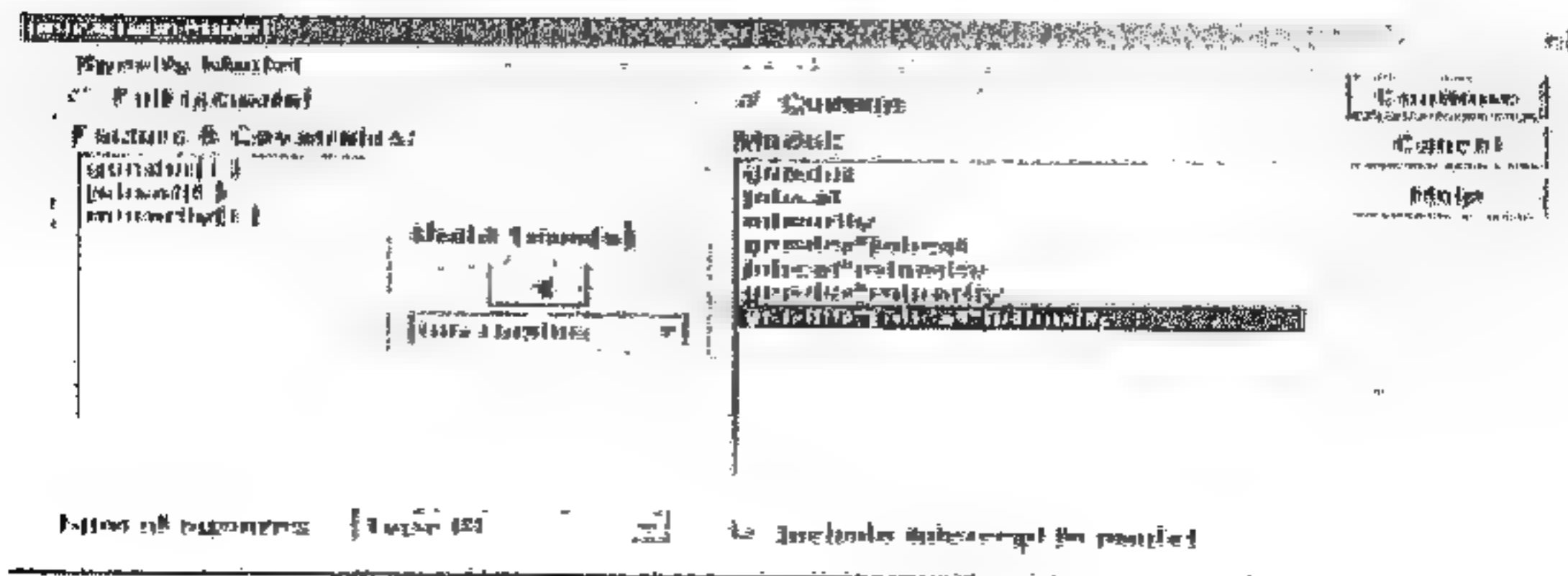
Univariate يظهر مربع الحوار التالي:



2. انقل المتغير Salary الى المستطيل اسفل Dependent Variable والمتغيرات

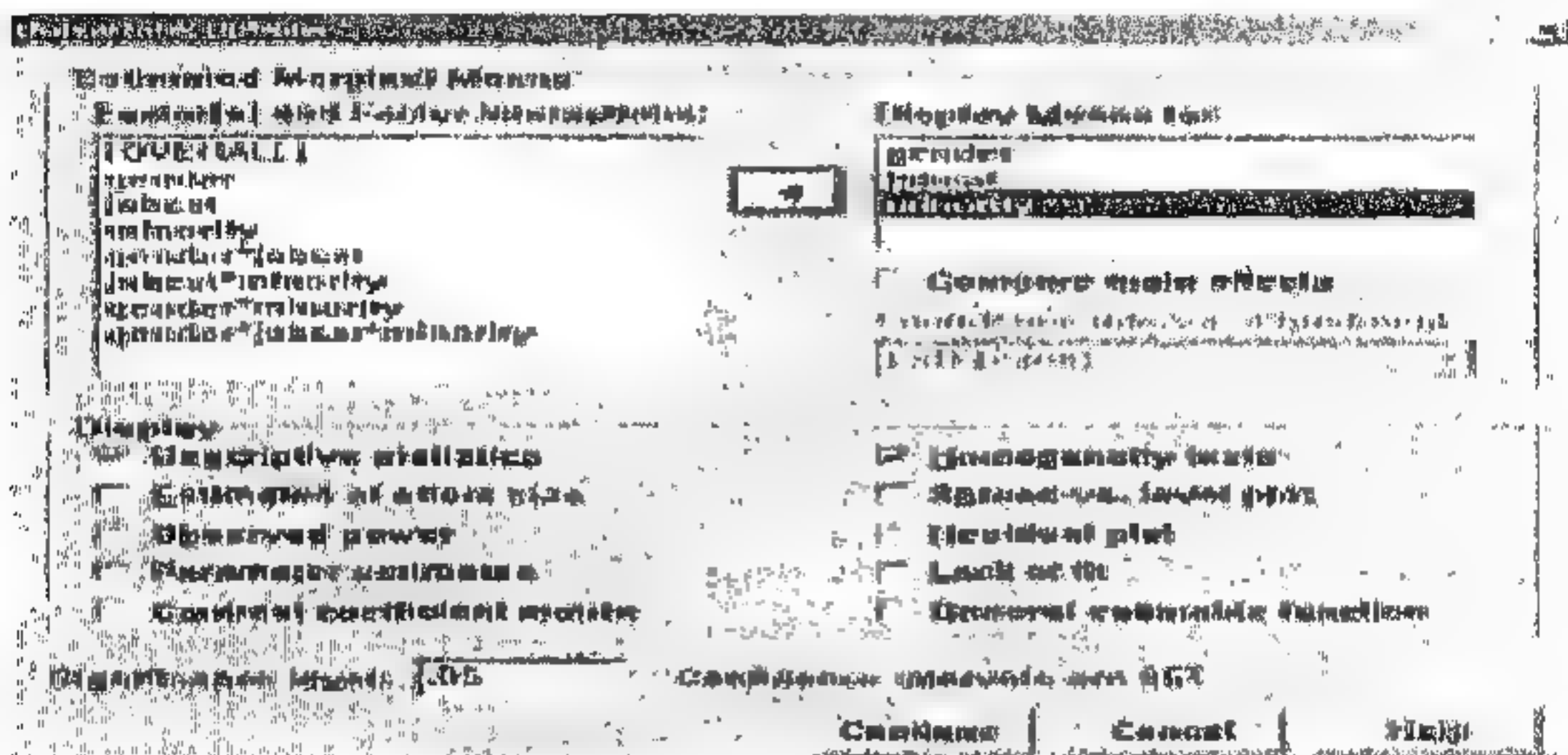
Gender و id و Jobcat الى المستطيل اسفل Fixed Factor(s) .

3. اضغط على Model يظهر مربع الحوار التالي:

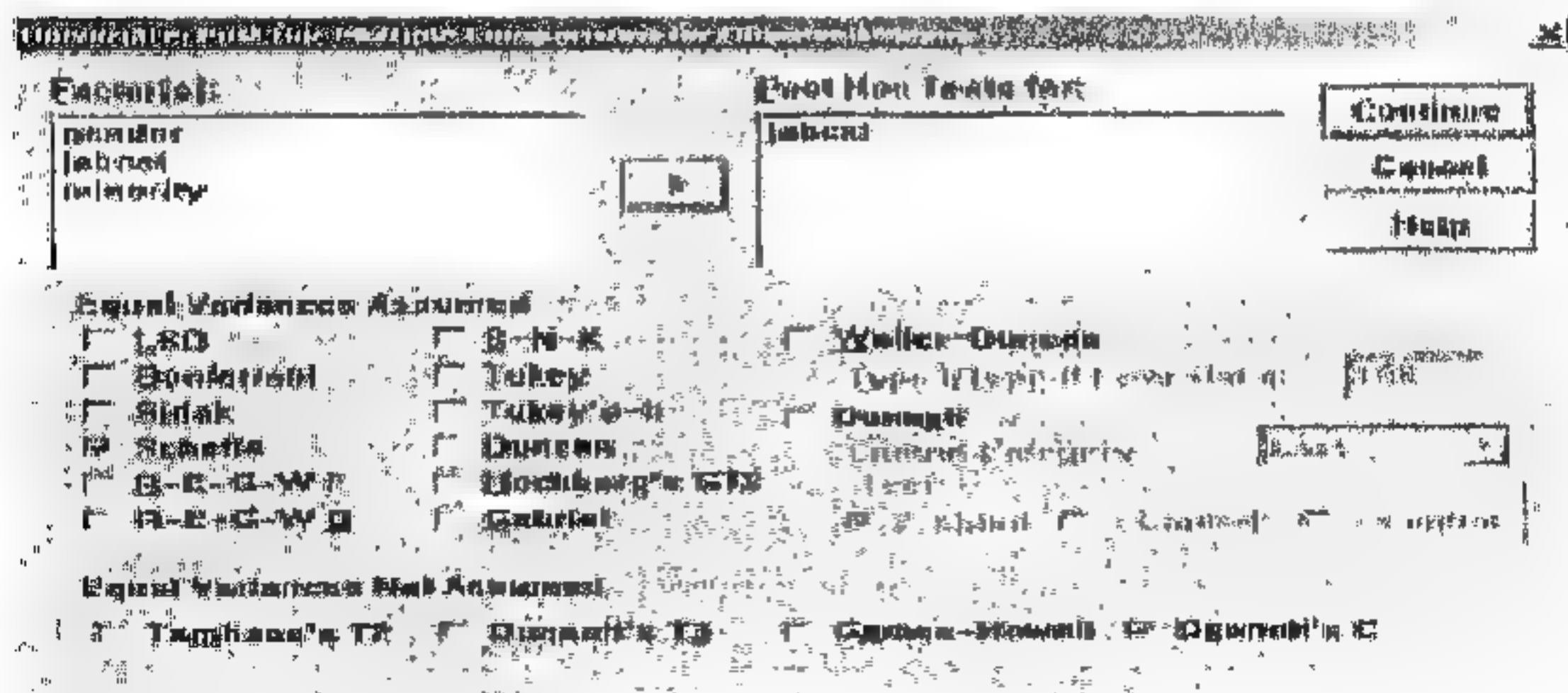


4. اختر Custom ثم Main effects من القائمة أسفل Build Term (s) وانقل المتعيرين Gender و Jobcat و id الى المستطيل أسفل Model ، ثم اختر Interaction من القائمة Build Term وانقل المتغيرات معا مثني مثني ثم جميعهم الى المستطيل أسفل Model كما بالشكل أعلاه، اضغط Continue سنعود إلى المربع الأصلي.

5. ضغط Options سيظهر مربع الحوار التالي اختر من Descriptive atatistics والخيار Homogeneity tests وانقل المتغيرات الثلاثة الى المستطيل أسفل Display Means for ثم اضغط Continue لنعود لمربع الحوار الأصلي.



6. اضغط على Post Hoc ليظهر مربع الحوار التالي:



7. اختر اختبار شفیه Scheffe للمقارنات البعدية من قائمة الاختبارات البعدية التي تشترط تماثل تباينات الفئات Equal Variance Assumed .

8. اختبر اختبار دونت س Dunnett,s C من قائمة الاختبارات البعدية التي لا تشترط

تمثل تباينات الفئات Equal Variance Not Assumed

9. انقل المتغير Jobcat فقط إلى المستطيل اسفل Post Hoc Tests For لأنه يتكون

من ثلاث مستويات أما المتغيران الاخران فلا ننقلهما لأنهما يتكونان من مستويين فقط.

10. اضغط Continue سنعود للمربع الحوار الأصلي. اضغط Ok تأج النتائج:

على الدارس تفسير النتائج

Univariate Analysis of Variance

Between-Subjects Factors

		Value Label	N
Minority Classification	0	No	370
	1	Yes	104
Gender	f	Female	216
	m	Male	258
Employment Category	1	Clerical	363
	2	Custodial	27
	3	Manager	84

Descriptive Statistics

Dependent Variable: Current Salary

Minority Classificat	Gender	Employment Cated	Mean	Std. Deviation	N
No	Female	Clerical	25,471.45	\$6,092.372	166
		Manager	47,213.50	\$8,501.253	10
		Total	26,706.79	\$8,011.894	176
	Male	Clerical	32,671.64	\$8,578.999	110
		Custodial	31,178.57	\$1,658.743	14
		Manager	55,683.57	\$18,029.451	70
		Total	44,475.41	\$20,330.662	194
	Total	Clerical	28,341.09	\$7,994.659	276
		Custodial	31,178.57	\$1,658.743	14
		Manager	53,374.81	\$18,164.043	80
		Total	86,023.31	\$18,044.096	370
Yes	Female	Clerical	23,062.50	\$3,972.369	40
		Total	23,062.50	\$3,972.369	40
	Male	Clerical	28,952.13	\$5,712.419	47
		Custodial	30,680.77	\$2,562.920	13
		Manager	76,037.50	\$17,821.961	4
		Total	32,246.09	\$13,059.881	64
	Total	Clerical	26,244.25	\$5,772.874	87
		Custodial	30,680.77	\$2,562.920	13
		Manager	76,037.50	\$17,821.961	4
		Total	28,713.94	\$11,421.638	104
Total	Female	Clerical	25,003.69	\$5,812.838	206
		Manager	47,213.50	\$8,501.253	10
		Total	26,031.92	\$7,558.021	216
	Male	Clerical	31,558.15	\$7,997.978	157
		Custodial	30,938.89	\$2,114.616	27
		Manager	56,243.24	\$18,051.570	74
		Total	41,441.78	\$19,499.214	258
	Total	Clerical	27,838.54	\$7,567.995	363
		Custodial	30,938.89	\$2,114.616	27
		Manager	53,977.80	\$18,244.776	84
		Total	84,419.57	\$17,075.661	474

Levene's Test of Equality of Error Variances^a

Dependent Variable: Current Salary

F	df1	df2	Sig.
17.696	8	465	.000

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design:

Intercept+MINORITY+GENDER+JOB CAT+MINORITY *

GENDER+GENDER * JOB CAT+MINORITY *

JOB CAT+MINORITY * GENDER * JOB CAT

Tests of Between-Subjects Effects

Dependent Variable: Current Salary

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	9.751E+10 ^a	8	1.219E+10	140.251	.000
Intercept	1.444E+11	1	1.444E+11	1661.526	.000
MINORITY	61989119.7	1	61989119.66	.713	.399
GENDER	756876310	1	756876310	54.737	.000
JOB CAT	2.006E+10	2	1.003E+10	115.420	.000
MINORITY * GENDER	27977363.9	1	27977363.93	.322	.571
GENDER * JOB CAT	981526336	1	981526335.9	11.294	.001
MINORITY * JOB CAT	690053398	2	345026699.0	3.970	.020
MINORITY * GENDER * JOB CAT	.000	0	.	.	.
Error	4.041E+10	465	86903667.84		
Total	6.995E+11	474			
Corrected Total	1.379E+11	473			

a. R Squared = .707 (Adjusted R Squared = .702)

Estimated Marginal Means

1. Minority Classification

Dependent Variable: Current Salary

Minority Classification	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
No	40443.745 ^a	835.531	38801.861	42085.629
Yes	39683.224 ^a	1423.737	36885.469	42480.979

a. Based on modified population marginal mean.

2. Employment Category

Dependent Variable: Current Salary

Employment Category	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Clerical	27539.427	577.449	26404.695	28674.160
Custodial	30929.670 ^a	1795.293	27401.779	34457.562
Manager	62978.190 ^a	1875.508	59292.670	66663.711

a. Based on modified population marginal mean.

Post Hoc Tests Employment Category

Multiple Comparisons

Dependent Variable: Current Salary

	(I) Employment Category	(J) Employment Category	Mean Difference (I-J)	Std. Error	Sig.
Scheffe	Clerical	Custodial	-\$3,100.35	\$1,859.586	.250
		Manager	-\$36,139.26*	\$1,128.703	.000
	Custodial	Clerical	\$3,100.35	\$1,859.586	.250
		Manager	-\$33,038.91*	\$2,062.334	.000
	Manager	Clerical	\$36,139.26*	\$1,128.703	.000
		Custodial	\$33,038.91*	\$2,062.334	.000
Dunnett C	Clerical	Custodial	-\$3,100.35*	\$568.679	
		Manager	-\$36,139.26*	\$2,029.912	
	Custodial	Clerical	\$3,100.35*	\$568.679	
		Manager	-\$33,038.91*	\$2,031.840	
	Manager	Clerical	\$36,139.26*	\$2,029.912	
		Custodial	\$33,038.91*	\$2,031.840	

Based on observed means.

*. The mean difference is significant at the .05 level.

4-9 تحليل التباين (ANCOVA) analysis of Covariance

يعتبر تحليل التباين (ANCOVA) من الأساليب المستخدمة مع التصميم شبه التجريبي والفرق بينهما أنه في حال التصميم التجريبي يتم ضبط المتغيرات المتحكم في المتغير التابع أما في التصميم شبه التجريبي فإن بعض المتغيرات المؤثرة في النتائج لم يتم إجراء تكافؤ بين المجموعات التجريبية والضابطة إنما تم قياس تلك المتغيرات فقط وبالتالي فإن استخدام الاختبار البديل وهو تحليل التباين لن يعطي نتائج حقيقية ، لذلك نستخدم اختبار تحليل التباين.

مثال: تم اختيار 9 مفحوصين عشوائياً و توزيعهم علي 3 مجموعات درست كل مجموعة منهم بإحدى طرق التدريس الرياضيات وتم قياس الاستعداد الرياضي والتحصيل حيث تم قياس الاستعداد الرياضي أو القدرة الرياضية قبل التجربة وقياس التحصيل بعد التجربة.

مجموعة 1		مجموعة 2		مجموعة 3	
استعداد	تحصيل	استعداد	تحصيل	استعداد	تحصيل
Apt	Ach	Apt	Ach	Apt	Ach
8	6	7	8	9	12
4	3	5	7	10	12
6	4	8	10	7	10

ونلاحظ أن البيانات يجب إدخالها بالترتيب التالي برنامج SPSS

الطريقة Method	Apt استعداد	Ach تحصيل
1	8	6
1	4	3
1	6	4
2	7	8
2	5	7
2	8	10
3	9	12
3	10	12
3	7	10

الآن لدينا ثلاثة متغيرات وهي:

● المتغير المستقل : طريقة بثلاث مستويات أو ثلاثة مجموعات مستقلة .

● متغير التباين : اختبار القدرة الرياضية .

● المتغير التابع : اختبار التحصيل في الرياضيات

والواقع أن اختبار انكوفيا يحاول حذف أثر المتغير المصاحب من المتغير التابع - التحصيل ومن ثم نستخدم اختبار - ف F علي المتوسطات المعدلة للاختبار التحصيلي لتبين الفروق بينها، وفي حال وجود أكثر من طريقتين أو عينتين يجب استخدام اختبارات المقارنة البعدية

وبرنامج SPSS لديه LSD و ت- بونفروني و Sidak المعدل لتحسس موضع الفروق . ولكن يفضل اختبار بونفروني و Sidak لأنهما يعتمدان علي المتوسطات المعدلة.

شكل البيانات بمحرر SPSS

	method	apt	ach
1	1.00	8.00	6.00
2	1.00	4.00	3.00
3	1.00	6.00	4.00
4	2.00	7.00	8.00
5	2.00	5.00	7.00
6	2.00	8.00	10.00
7	3.00	9.00	12.00
8	3.00	10.00	12.00
9	3.00	7.00	10.00

التحقق من فرضية تساوي ميل الانحدار :

حتى نصبح تحليل التغير يجب أن تكون العلاقة بين متغير التغير والمتغير التابع خطية وعدم وجود تفاعل بين متغير التغير والمعالجات

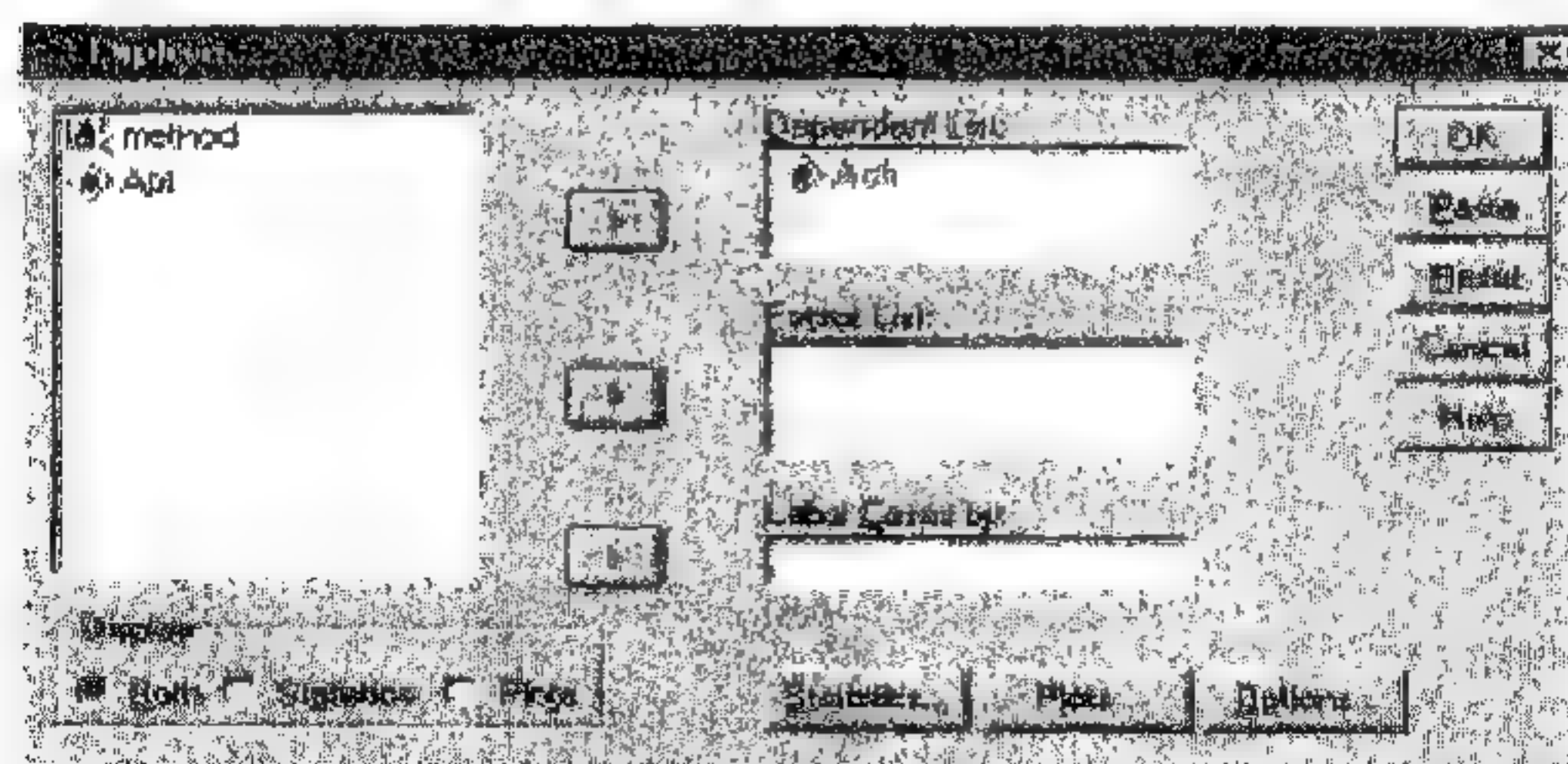
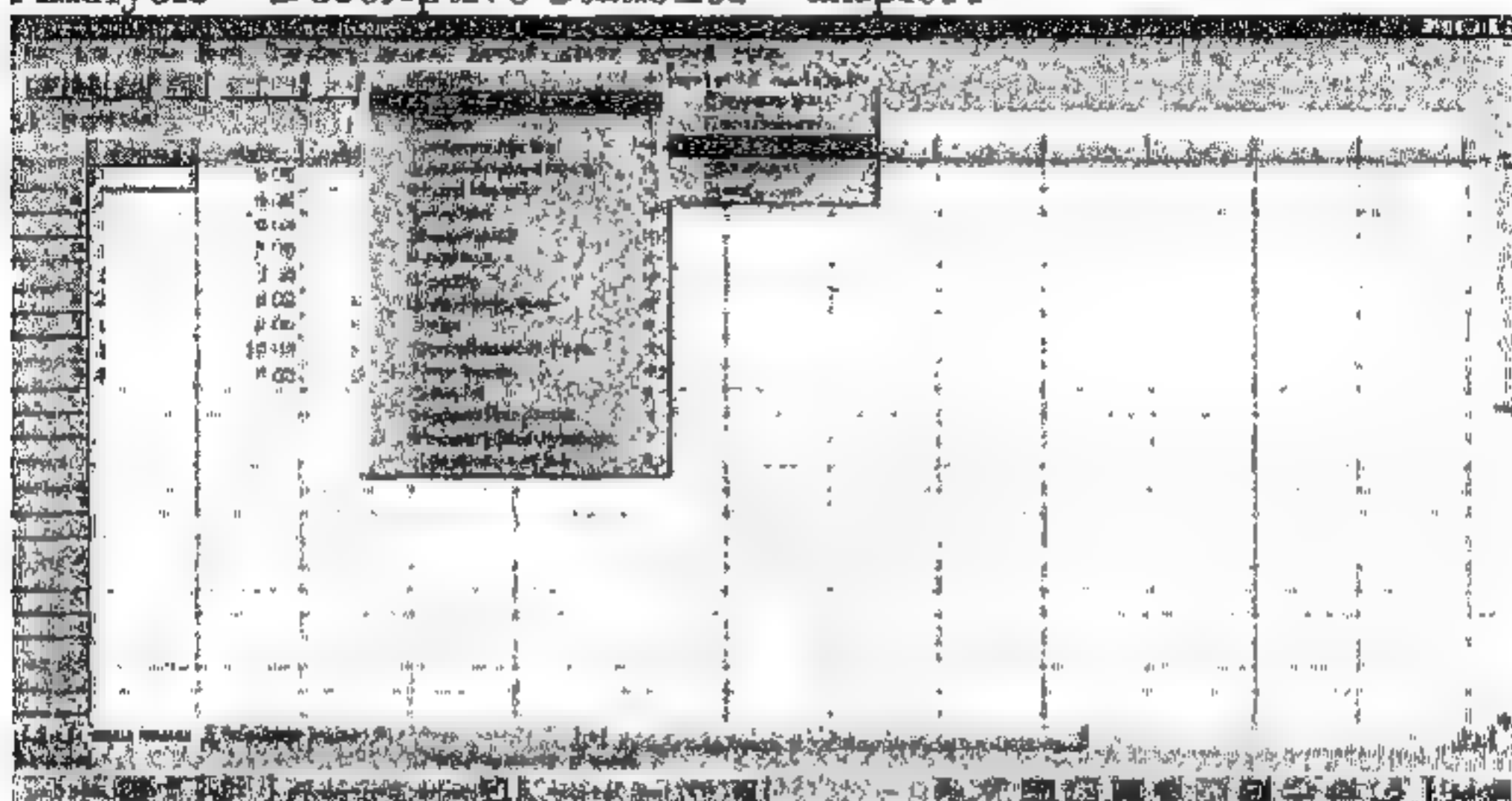
هل توجد علاقة خطية قوية بين متغير التغير و المتغير التابع؟ هل ميل الخطوط لكل من متغير التغير - الاستعداد - ومتغير التحصيل في المجموعات الثلاث متساوي ؟ (هل التفاعل بين الطرق الثلاث و الاستعداد دال إحصائياً؟

§ الخطية : نظراً لأن اختبار أنكرفا عبارة عن نموذج خطي عام مع المحددات متعددة ، فيفترض أن متغير التغير يرتبط خطياً مع المتغير التابع ، وفي مثالنا يجب أن يكون الارتباط خطي بين كل من الاستعداد الرياضي والتحصيل في الرياضيات ، وفي برنامج SPSS يجب أن نرسم أو نحدد الانحدار الخطي لكل المجموعات بين المتغيرين ثم نحسب مربع معامل الارتباط ونقارن بين الميل في كل مجموعة ..

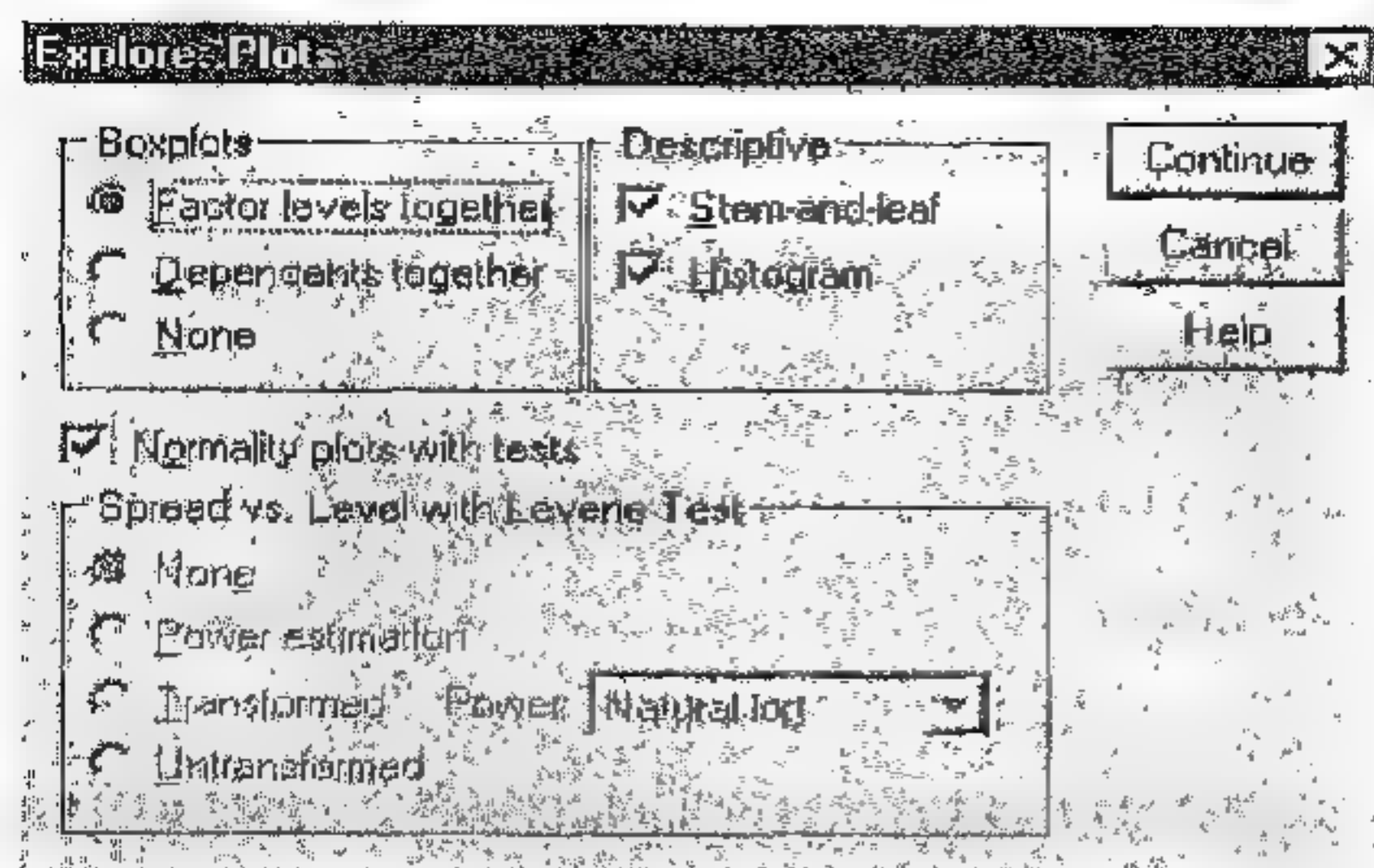
§ تجانس التباين : يجب أن يكون التباين متجانساً لكل المجموعات بين المتغيرين وهو نفس شرط اختبار أنوفا مع ملاحظة شرط الاعتدالية أيضاً ، و شرط الاعتدالية شرط رياضي مهم ، وقد اقترح بعض الباحثين أن هذا الشرط يمكن التغاضي عنه في حال العينات التي تزيد

عس 30 وفي مثالنا العينة 3 ولكن فؤاد أبو حطب نقلا عن باحثين آخرين ذكر بتقليص العدد لـ 15 علي أساس أن الفروق تكون ضئيلة ، ولكن برنامج مثل sigmaStat يقوم بإجراء هذه العملية في كل مرة وللأسف برنامج SPSS يترك ذلك لخبرة مشغل البرنامج الذي عليه ان يحقق من شرطي التجانس وشرط الاعتدالية بنفسه باستخدام اختبار ليفين لتجانس التباين واختبار ليفور أو اختبار Shapiro-Wilk أو اختبار كلمجروف - سميرونوف Kolmogorov-Smirnov ونستطيع إجراهما من خلال برنامج SPSS وفق الخطوات التالية :

Analysis - Descriptive Statistics - Explore



انقر الزر Plots تحصل علي النافذة التالية:



عدم خانة Normality plots with tests وانقر الزر Continue تظهر النافذة التي تسبقها ثم انقر موافق .

Tests of Normality						
Shapiro-Wilk			Kolmogorov-Smirnov(a)			
Sig.	Df	Statistic	Sig.	df	Statistic	
.537	9	.936	.200(*)	9	.174	Ach

* This is a lower bound of the true significance.

a Lilliefors Significance Correction

ويخشى بعض الباحثين عدم تحقق شرط الاعتدالية والحل ببساطة هو اللجوء للإحصاء الابارامترك.

§ تجانس الانحدار : نفترض أنكوفا تجانس الانحدار وبعض البرامج تحسبه بشكل مستقل ولكن يستندل عليه في برنامج SPSS بحساب دلالة التفاعل بين المتغيرين من خلال حساب جدول أنكوفا مع حساب التفاعل.

حساب تجانس الانحدار مع باستخدام برنامج SPSS

1. اختر Analysis - General Linear Model - Univariate.

§ اختر المتغير ach وضعه في مربع المتغير التابع Independent variable .

▪ اختر method وضعها في مربع العامل الثابت Fixed Factor

▪ اختر المتغير apt وضعه في مربع متغير التفاعل Covariate

METHOD*APT.

Tests of Between-Subjects Effects

Dependent Variable: ACH

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Eta Squared
Corrected Model	24.605 ^a	5	4.921	40.913	.005	.987
Intercept	1.629	1	1.629	4.461	.125	.598
METHOD	1.416	2	.708	1.938	.289	.564
APT	10.343	1	10.343	28.332	.013	.904
METHOD*APT	.128	2	6.456E-02	.177	.846	.105
Error	1.095	3	.365			
Total	662.000	9				
Corrected Total	58.000	8				

R Squared = .987 Adjusted R Squared = .986

$$Partial \eta^2_{interaction} = \frac{Sum of Square_{interaction}}{Sum of Square_{interaction} + Sum of Square_{error}}$$

الاستخلاص:

The interaction is not significant, $F(2,3) = .177$, $p = .846$, partial eta square = .105.

التفاعل ليس دال إحصائياً مما يشير لتحقيق شرط تجانس الانحدار وبذلك يتحقق أحد

شروط تحليل التباين

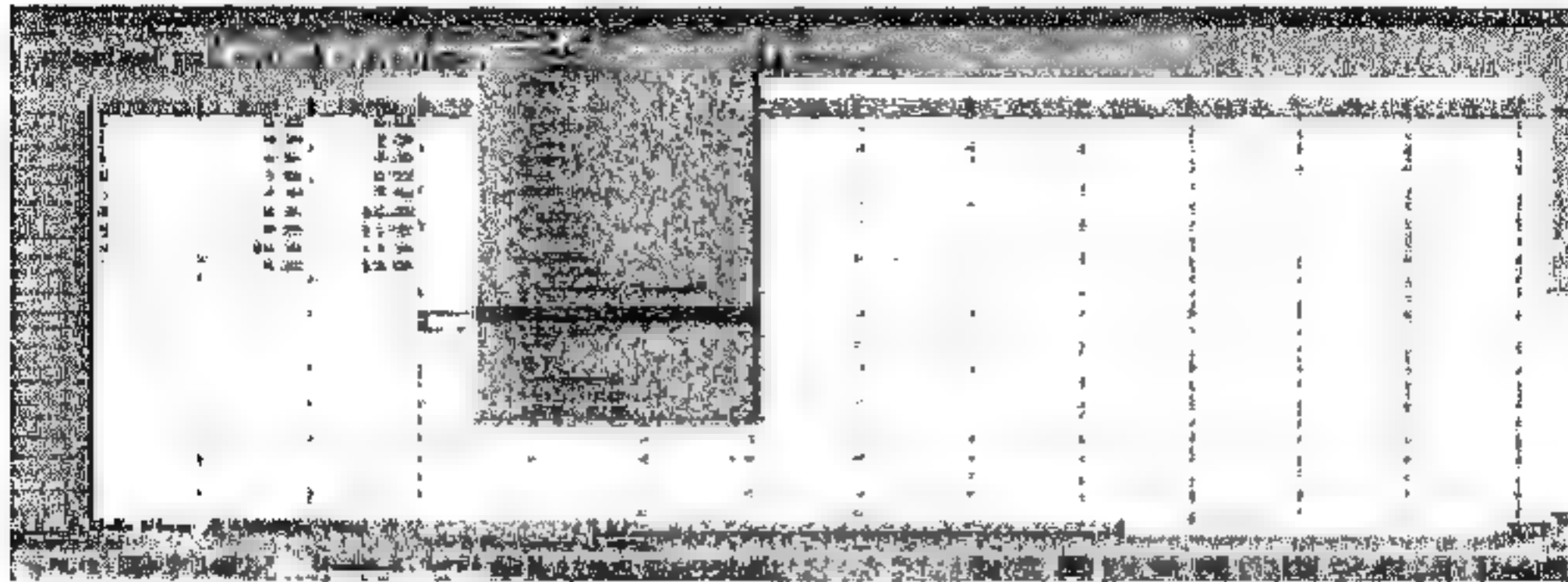
وبعض البرامج الأخرى غير برنامج SPSS تخرج جدول مستقل لتجانس الانحدار كما

يلي :

اختبار تجانس الانحدار

Source	SS	df	MS	F	P
between regressions	0.13	2	0.06	0.18	0.843671
remainder	1.1	3	0.37		
adjusted error	1.22	5			

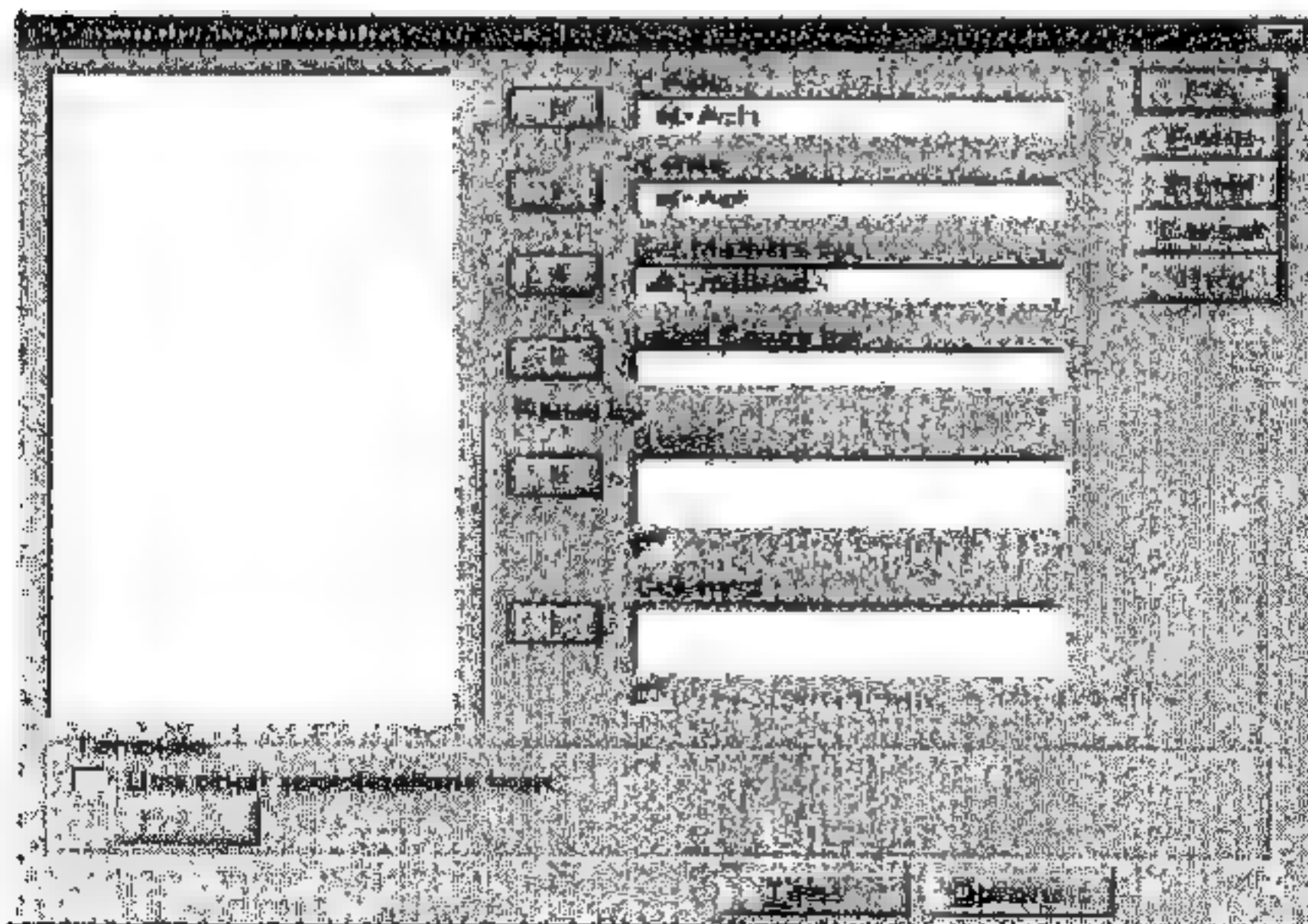
وهذا الجدول لا يظهر في برنامج SPSS وعدم دلالة قيمة F تعني تجانس الانحدار
 اختبار شرط العلاقة الخطية بين المتغير التابع - التحصيل - ومتغير التباين الاستعداد
 من قائمة رسم Graph اختر Scatter ثم اختر النمط النشط البسيط



اختيار التشتت البسيط

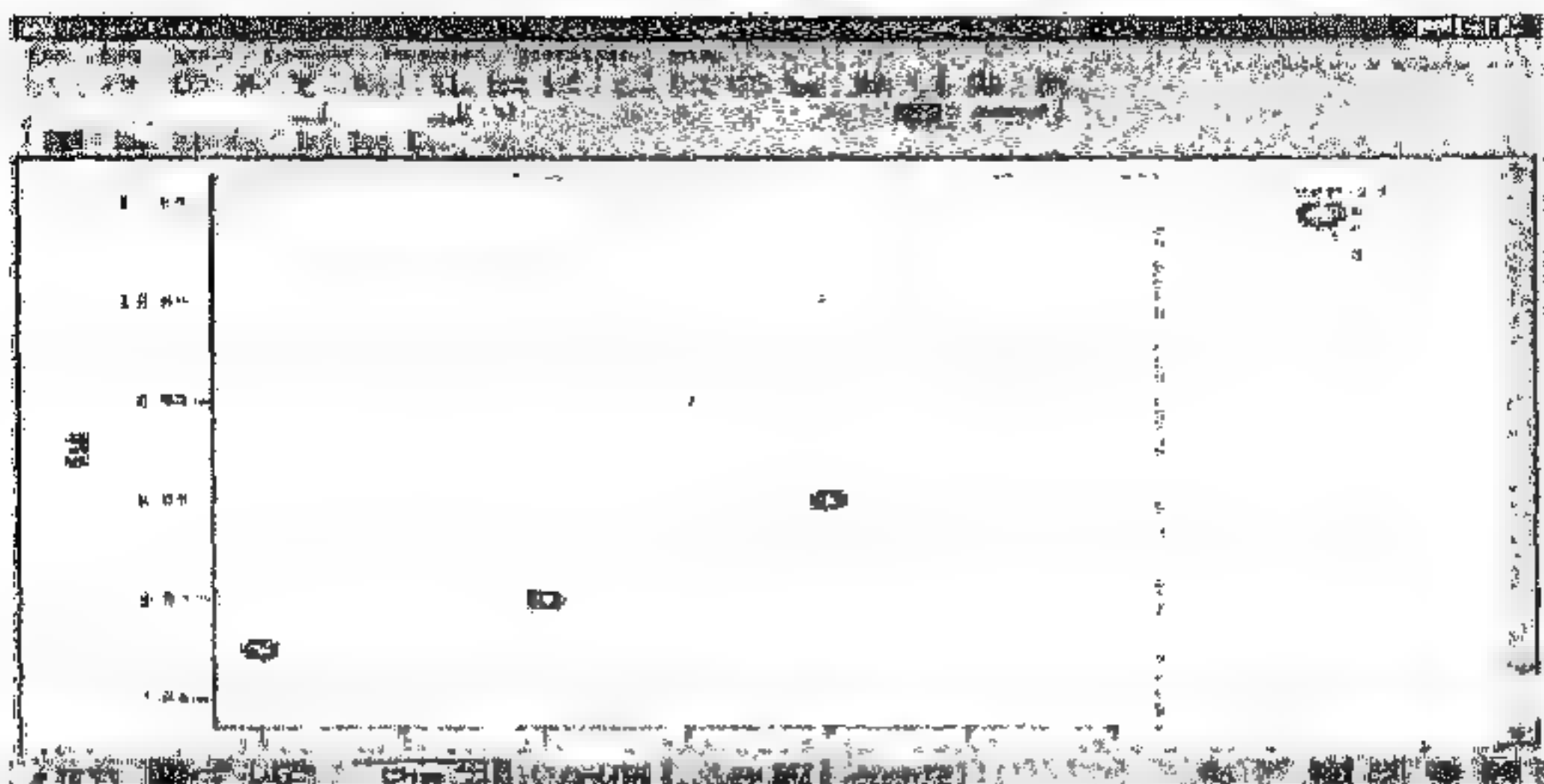


انقر الزر Define وانقل المتغير Ach لمربع محور Y ثم ضع المتغير Apt في مربع محور X
 وانقل متغير العامل method لمربع Set Markers by كما بالشكل التالي :
 نافذة اعدادات رسم الانحدار البسيط

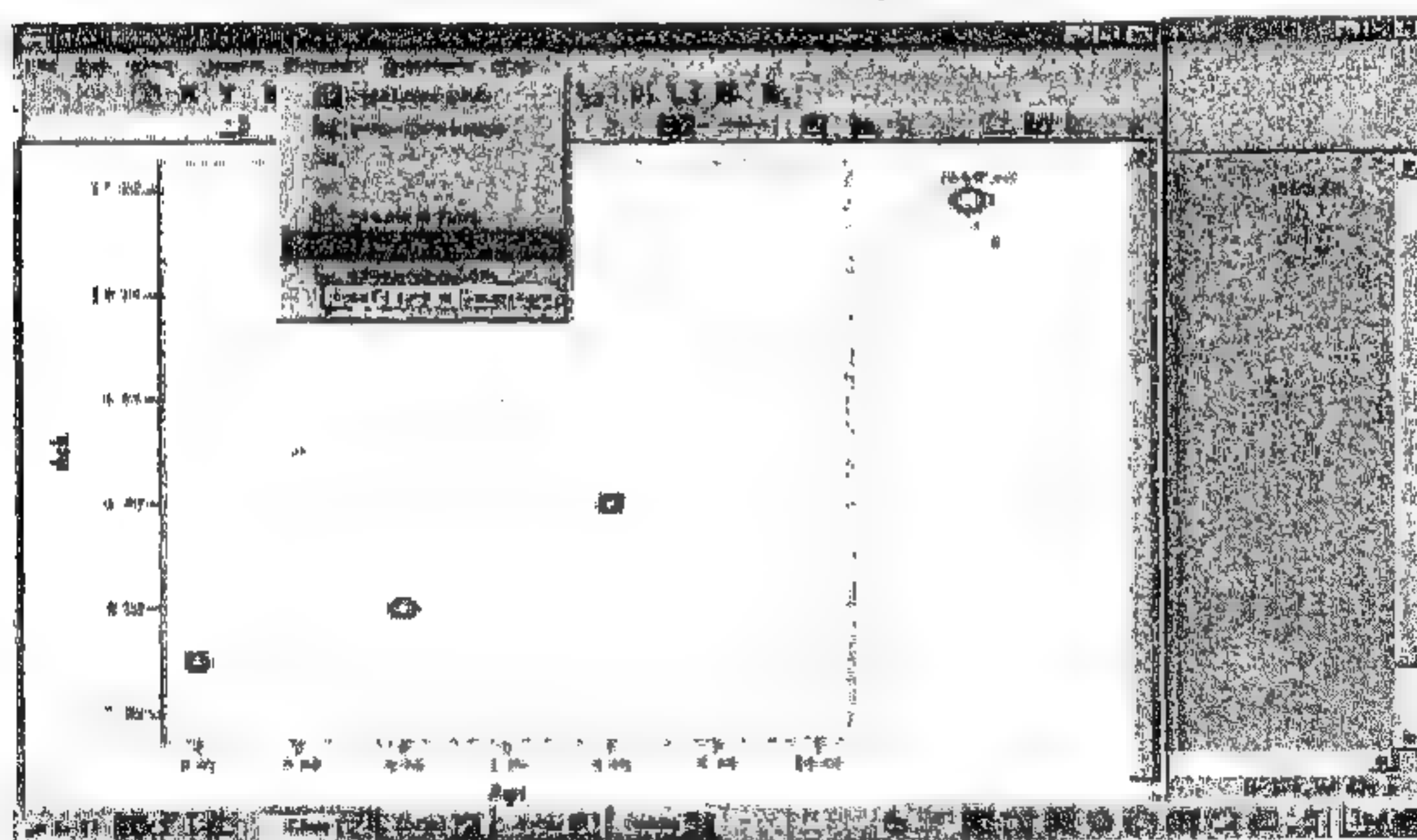


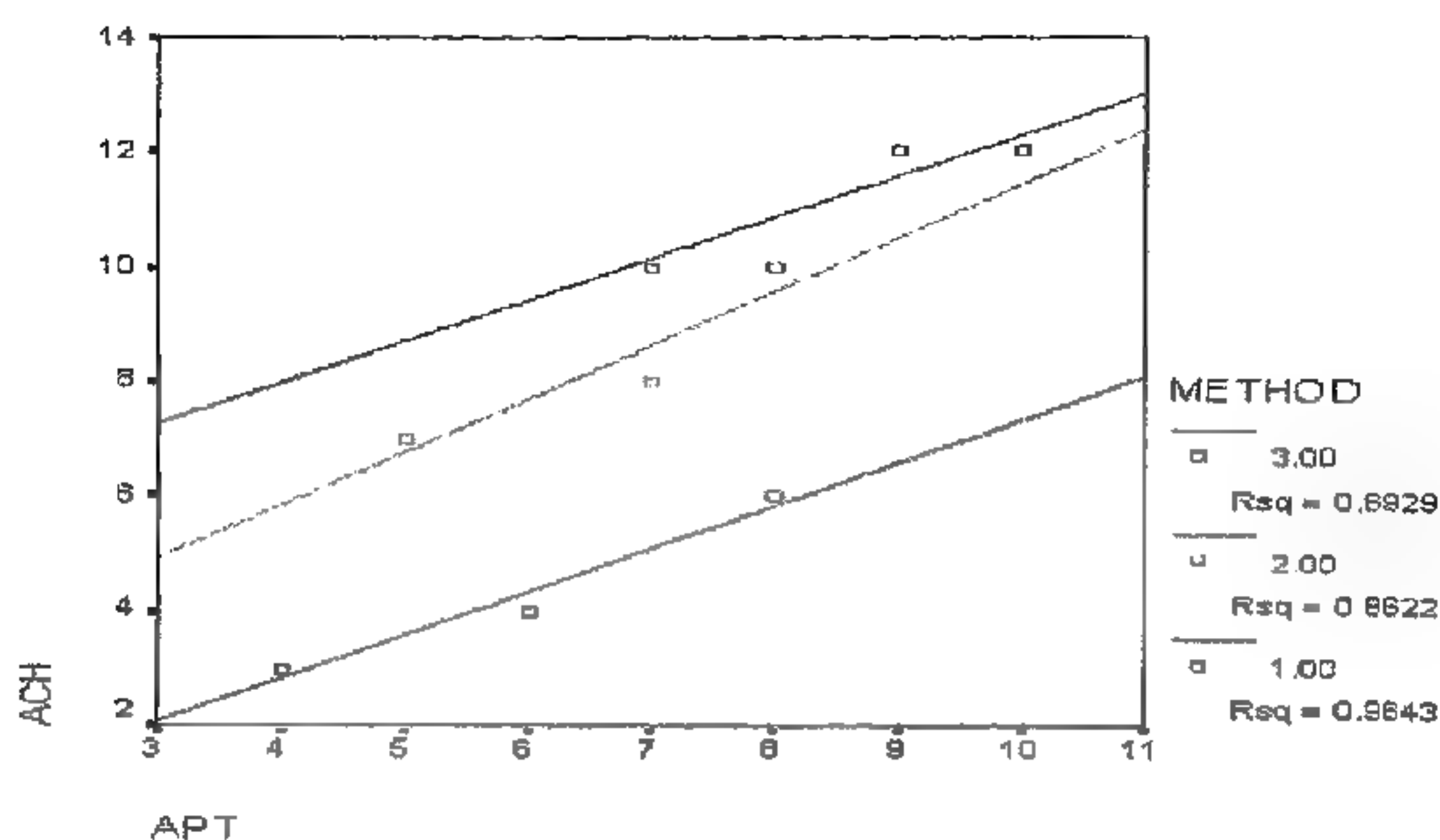
عدل المخطط للحصول علي خطوط الانحدار بنقر الرسم نقرأ مزدوجاً

انقر دائرة الطريقة الأولى ومن ثم تعلم نقاطها



من قائمة elements اختر Fit subgroup





ومن الرسم يتضح أن متغير التغير-الاستعداد الرياضي - يرتبط بعلاقة خطية مع المتغير التابع - التحصيل في الرياضيات - في الطرق الثلاث المستخدمة في التدريس ومن ثم يتحقق شرط تساوي الميل The assumption of equal slopes ومن ثم ندرى اختبار تحليل التغير أنكوفاً باتجاه واحد .

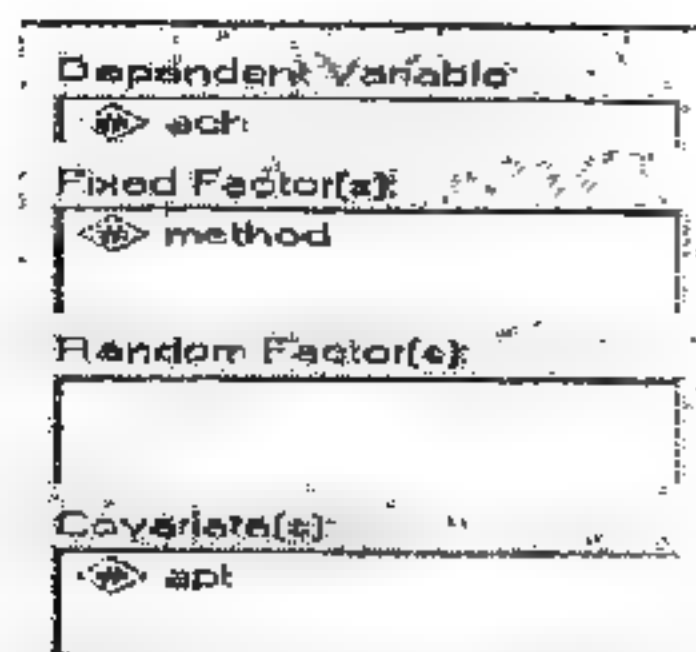
وهذا واضح من شبه التوازي في الطرق الثلاث و مربع معامل الارتباط مرتفع والمتقارب في القيمة

تطبيق اختبار أنكوفاً

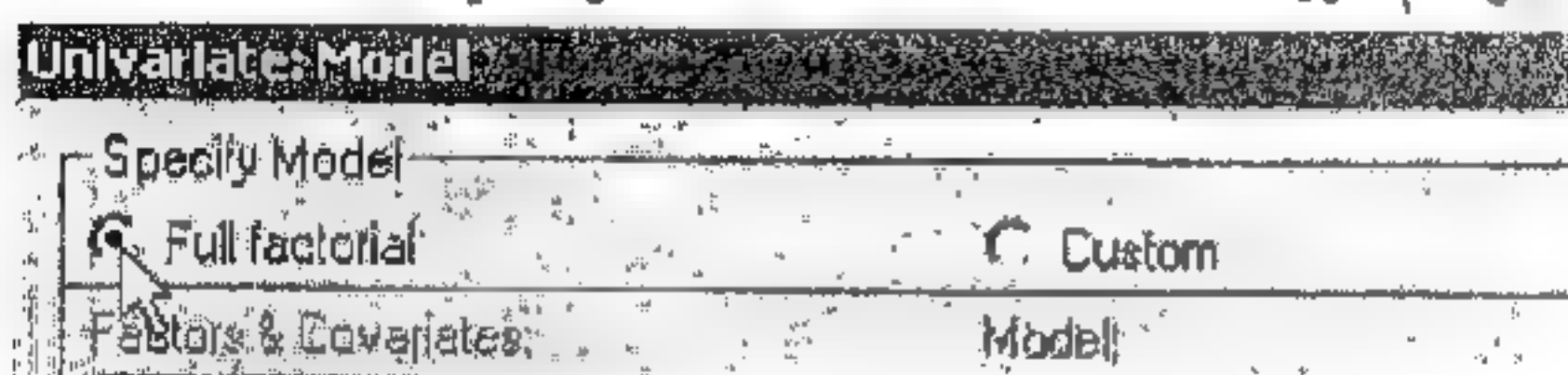
الفروق في التحصيل لا ترجع فقط للفروق عبر طرق التدريس الثلاث ولكن أيضاً للفروق الابتدائية في الاستعداد ولكي نزيل أثر الاستعداد كما نراها في التطبيق القبلي لاختبار الاستعداد الرياضي ،ومن ثم نستخدم اختبار أنكوفاً الذي سوف نناقشه فيما يلي من خلال أمثال ولكن أو التنويه إلى أن الخطوات السابقة ضرورية ويجب التوقف عن استخدام اختبار أنكوفاً لدي عدم تحقق تلك الشروط وخاصة شرطي تجانس الانحدار و شرط الخطية الذي يشير ببساطة لارتباط تنبؤي بين متغير التغير والمتغير التابع . وسوف نفترض أنك أجريت تلك الاختبارات التحقيقية:

اختر Analyze \ General Linear Model \ Univariate وضع المتغيرات كما سبق

ويوضحها الشكل التالي



اختر Model وعلم الزر Full factorial كما بالشكل التالي :



انقر الزر Ok

المتوسطات المعدلة

Estimated Marginal Means

METHOD

Dependent Variable: ACH

METHOD	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1.00	5.209 ^a	.315	4.400	6.018
2.00	8.684 ^a	.291	7.937	9.431
3.00	10.107 ^a	.340	9.232	10.982

a.

Evaluated at covariates appeared in the model: APT = 7.1111.

الآن يجب أن نختبر الفروق بين المتوسطات لنحسب الفرق بين الطرق الثلاث بعد عزل

أثر الاستعداد الرياضي

Tests of Between Subjects Effects

Dependent Variable: ACH

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Eta Squared
Corrected Model	84.776 ^a	3	28.259	115.401	.000	.986
Intercept	1.891	1	1.891	7.723	.039	.607
APT	10.776	1	10.776	44.005	.001	.898
METHOD	26.587	2	13.294	54.288	.000	.956
Error	1.224	5	.245			
Total	86.000	9				
Corrected Total	86.000	8				

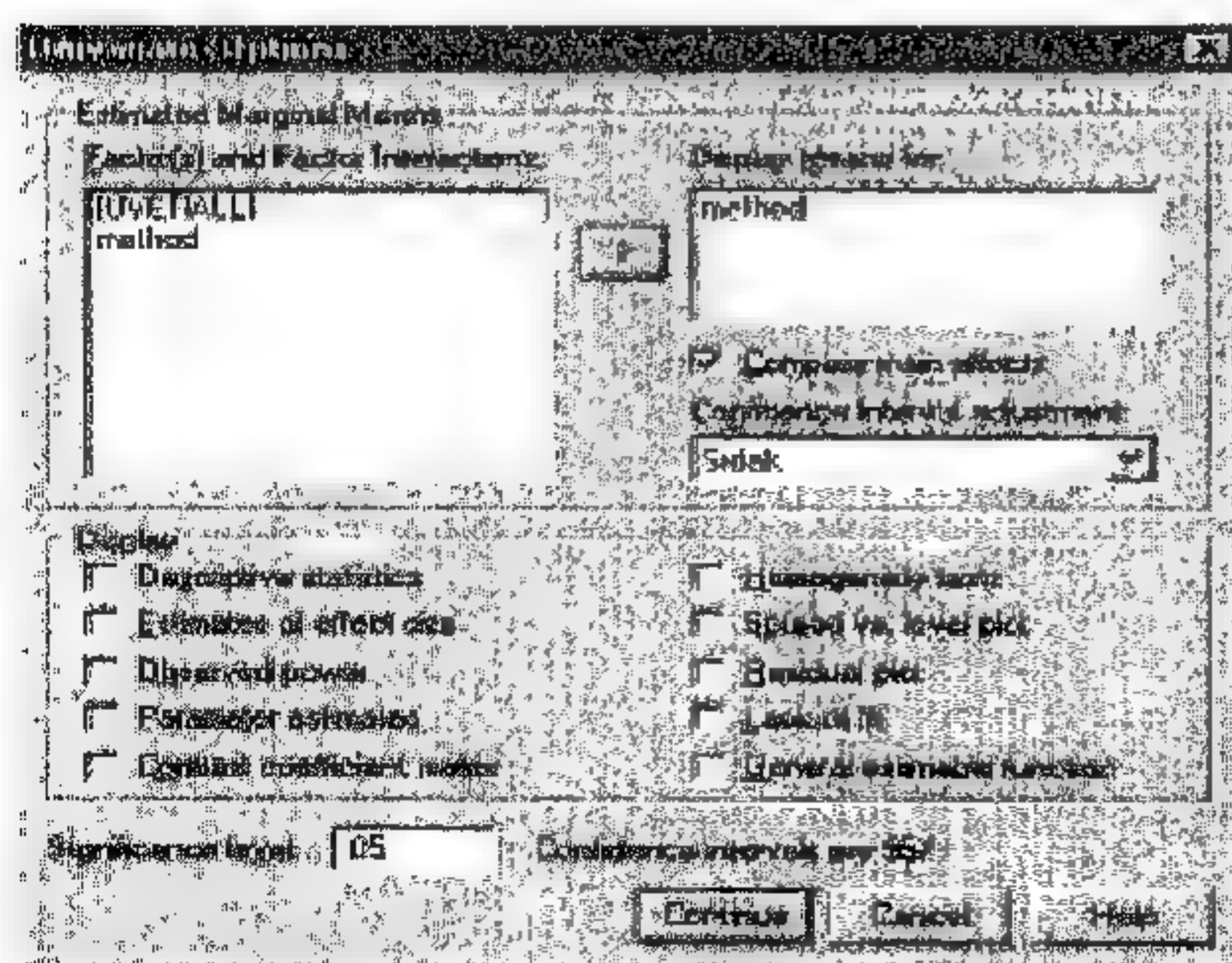
a. R Squared = .986 (Adjusted R Squared = .977)

من الجدول ننظر لقيمة F المقابلة للطرق المستخدمة Method نجد أن قيمة F 54.288 وهي دالة عند مستوى أقل من 0.001 وبالتالي يجب أن نستخدم اختبار للمقارنة البعدية لتحديد أي الطرق تفوقت علي الطريقتين الأخرتين ، كما نلاحظ أن حجم التأثير كبير 0.956

$$Partial \eta^2_{method} = \frac{Sum of Square_{method}}{Sum of Square_{method} + Sum of Square_{error}}$$

ملحوظة: لن نتطرق لاستخدام الزر contrast للسهولة ولذلك من نافذة الخيارات

السابقة



نتأكد من تعليم مربع الاختيار **compare main effect** وسوف نحصل على نتائج من ضمنها الجدول التالي:

المقارنات البعدية باستخدام اختبار بونفروني

Pairwise Comparisons

Dependent Variable: Ach

(I) method	(J) method	Mean Difference (I-J)	Std. Error	Sig. ^a	95% Confidence Interval for Difference ^a	
					Lower Bound	Upper Bound
1	2	-3.474*	.412	.001	-4.929	-2.019
	3	-4.897*	.514	.001	-6.712	-3.083
2	1	3.474*	.412	.001	2.019	4.929
	3	-1.423	.469	.087	-3.080	.234
3	1	4.897*	.514	.001	3.083	6.712
	2	1.423	.469	.087	-.234	3.080

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Bonferroni.

1- لا توجد فروق دالة إحصائية عند مستوى 0.01 بين الطريقة 2 والطريقة 3 لصالح الطريقة 2 ونلاحظ ظهور علامة * مع المتوسطين

2- توجد فروق دالة إحصائية عند مستوى 0.01 بين الطريقة 1 والطريقة 3 لصالح الطريقة 1

3- توجد فروق دالة إحصائية عند مستوى 0.01 بين الطريقة 1 والطريقة 2 لصالح الطريقة 1

ملحوظة: اختبار أنكوفنا بديل لاختبار أنوفنا في حالة عدم تكافؤ المجموعات.

5-9 استخدام برمجية SPSS في حل المسائل.

مثال 9-1: يوجد 4 شعب في كل شعبة 10 طلاب درست بأربع طرق مختلفة ، والمطلوب معرفة إن كان هناك فرق معنوي في متوسط الشعب الأربعة عند مستوى معنوية 0.10

الفرضية:

هل تختلف متوسطات الشعب باختلاف طريقة التدريس عند مستوى معنوية 0.10 ؟
المتغيرات:

المتغير التابع Dependent : العلامة Mark

المتغير المستقل Independent : الشعبة Sectionn ولها أربع مستويات.

إذاً نستخدم تحليل التباين الاحادي One Way ANOVA

إدخال البيانات:

	sec	mark		sec	mark
1	1.00	5.00	21	3.00	7.00
2	1.00	6.00	22	3.00	5.00
3	1.00	3.00	23	3.00	6.00
4	1.00	2.00	24	3.00	8.00
5	1.00	4.00	25	3.00	5.00
6	1.00	10.00	26	3.00	10.00
7	1.00	7.00	27	3.00	7.00
8	1.00	3.00	28	3.00	3.00
9	1.00	4.00	29	3.00	4.00
10	1.00	7.00	30	3.00	6.00
11	2.00	8.00	31	4.00	10.00
12	2.00	7.00	32	4.00	8.00
13	2.00	7.00	33	4.00	9.00
14	2.00	9.00	34	4.00	9.00
15	2.00	2.00	35	4.00	4.00
16	2.00	9.00	36	4.00	9.00
17	2.00	8.00	37	4.00	9.00
18	2.00	4.00	38	4.00	5.00
19	2.00	5.00	39	4.00	6.00
20	2.00	8.00	40	4.00	7.00

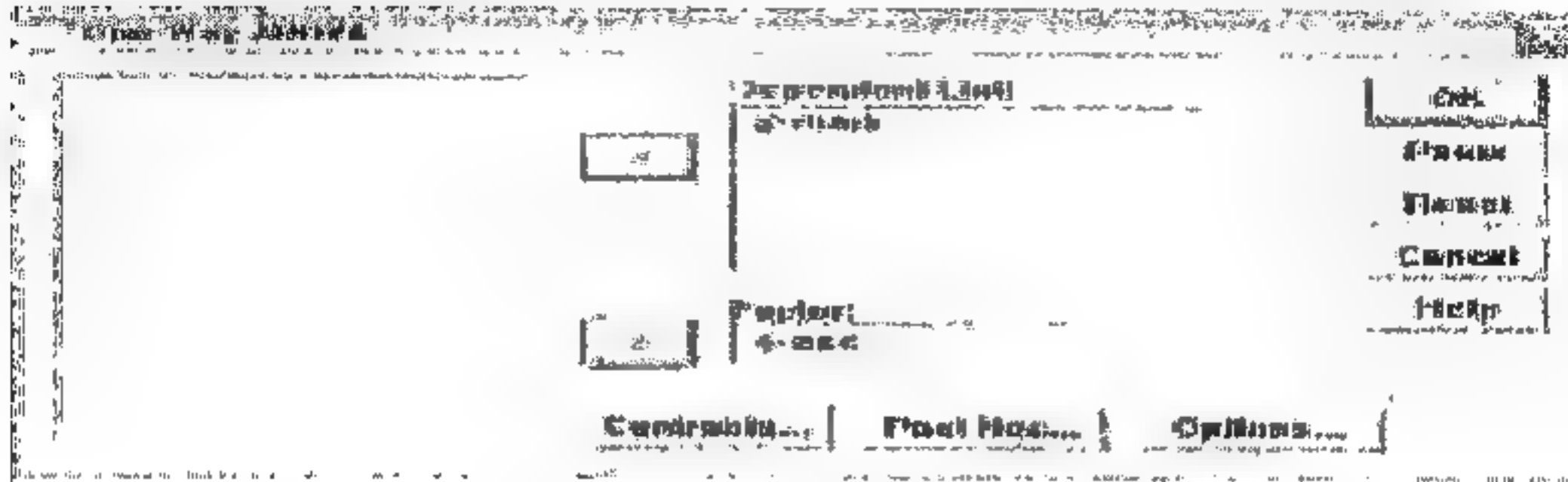
يستخدم اختبار ليفين للتأكد من تماثل التباينات Leven's homogeneity of variances test

لإجراء تحليل التباين تتبع الخطوات التالية:

Analyze - Compare Means - One-Way ANOVA...

ضع المتغير Mark في نافذة المتغيرات التابعة Dependent List:

ضع المتغير العامل sec في نافذة عاملي Factor:

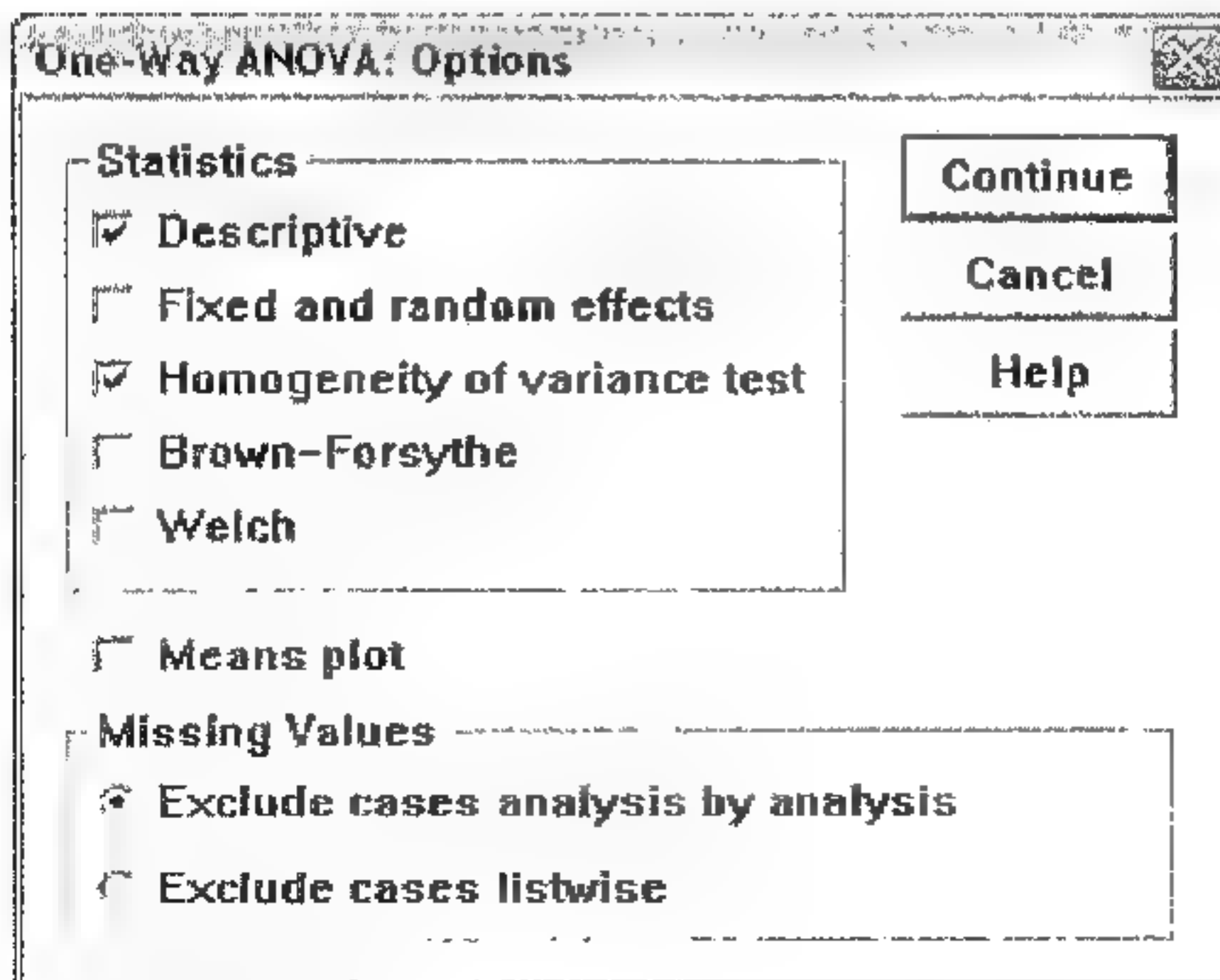


انقر مفتاح Options... تظهر لك شاشة الحوار أدناه

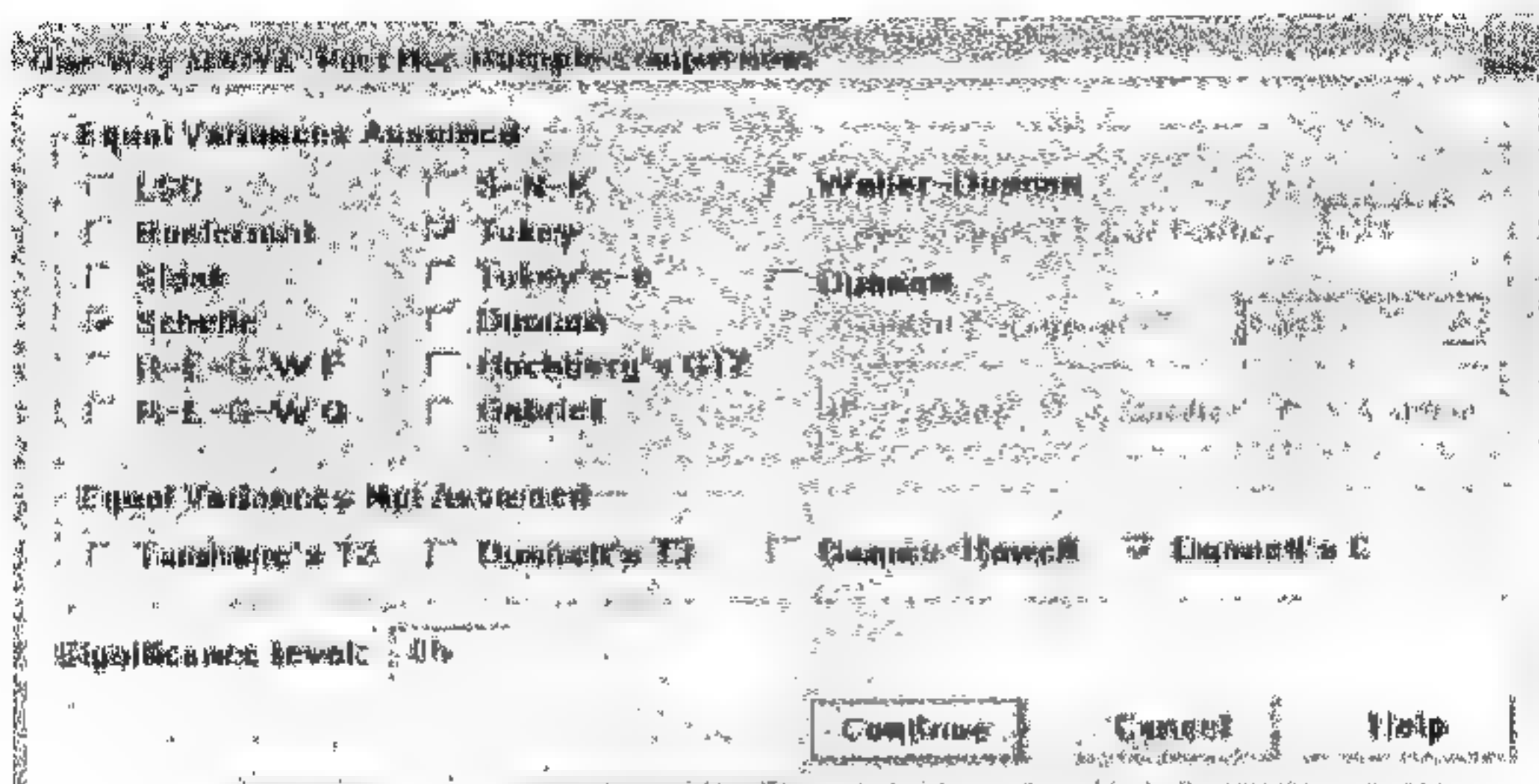
اختر حساب الإحصاءات الوصفية Descriptive

اختر فحص التماثل بين المجموعات Homogeneity of variance test

ثم انقر زر Continue فتعود الى شاشة حوار One-Way ANOVA



Post Hoc الاختبارات البعدية



من الاختبارات البعدية في حالة تساوي التباين اختراي Tukey و Scheffe
 من الاختبارات البعدية في حالة عدم تساوي التباين اختراي Dunnett's C
 ثم انقر زر Continue فتعود الى شاشة حوار One-Way ANOVA ثم انقر زر Ok
 تظهر نتيجة تحليل التباين الأحادي كما هو مبين أدناه:
 الإحصاءات الوصفية للمتغير التابع لكل فئة من فئات المتغير العامل sec

Oneway

Descriptives

MAF:

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1.00	10	5.1000	2.42441	.76667	3.3657	6.8343	2.00	10.00
2.00	10	6.7000	2.31181	.73106	5.0462	8.3538	2.00	9.00
3.00	10	6.1000	2.02485	.64031	4.6515	7.5485	3.00	10.00
4.00	10	7.6000	2.01108	.63596	6.1614	9.0386	4.00	10.00
Total	40	6.3750	2.30570	.36400	5.6375	7.1125	2.00	10.00

نتائج تحليل التباين الأحادي: الإحصاءات الوصفية للمتغير التابع لكل فئة من فئات المتغير
 عاملي مثل المتوسطات الحسابية Mean، الانحرافات المعيارية Std. Deviation، وأقصى المعاري
 Std. Error، وفترات الثقة 95% Confidence Interval for Mean، وأقل قيمة
 Minimum واكبر قيمة Maximum للمتغير التابع لكل فئة من فئات المتغير العاملي.

Test of Homogeneity of Variance

MARK

Levene Statistic	df1	df2	Sig.
.205	3	36	.892

نتائج تحليل التباين الاحادي: اختبار ليفين لفحص تجانس التباين لفئات المتغير العامي
ينتج أن تباين المجموعات متساوية لأن قيمة مستوى الدلالة Sig=0.892 وهي اكبر من مستوى
الدلالة ($\alpha = 0.05$).

ANOVA

MARK

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	33.075	3	11.025	2.277	.085
Within Groups	174.300	36	4.842		
Total	207.375	39			

نتائج تحليل التباين الأحادي: فحص فرضية الدراسة.
عدم وجود فروق ذات دلالة احصائية على مستوى أقل من ($\alpha = 0.05$)، حيث كانت
قيمة مستوى الدلالة Sig أقل من 0.05

Post Hoc Tests اختبارات الفروقات البعدية Multiple Comparisons

Dependent Variable: MARK

	(I) SEC	(J) SEC	Mean Difference (I-J)	Std. Error	Sig.	95 % Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	1.00	2.00	-1.6000	.98404	.377	-4.2602	1.0602
		3.00	-1.0000	.98404	.741	-3.6602	1.6602
		4.00	-2.5000	.98404	.070	-5.1602	.1502
	2.00	1.00	1.6000	.98404	.377	-1.0602	4.2602
		3.00	.6000	.98404	.928	-2.0602	3.2602
		4.00	-.9000	.98404	.797	-3.5602	1.7502
	3.00	1.00	1.0000	.98404	.741	-1.6602	3.6602
		2.00	-.6000	.98404	.928	-3.2602	2.0602
		4.00	-1.6000	.98404	.434	-4.1602	1.1502
	4.00	1.00	2.5000	.98404	.070	-.1502	5.1602
		2.00	.9000	.98404	.797	-1.7602	3.6502
		3.00	1.5000	.98404	.434	-1.1502	4.1502
Scheffe	1.00	2.00	-1.6000	.98404	.460	-4.4856	1.2856
		3.00	-1.0000	.98404	.793	-3.8856	1.8856
		4.00	-2.5000	.98404	.111	-5.3856	.3856
	2.00	1.00	1.6000	.98404	.460	-1.2856	4.4856
		3.00	.6000	.98404	.946	-2.2856	3.4856
		4.00	-.9000	.98404	.840	-3.7856	1.9856
	3.00	1.00	1.0000	.98404	.793	-1.8856	3.8856
		2.00	-.6000	.98404	.946	-3.4856	2.2856
		4.00	-1.6000	.98404	.516	-4.3856	1.3856
	4.00	1.00	2.5000	.98404	.111	-.3856	5.3856
		2.00	.9000	.98404	.840	-1.9856	3.7856
		3.00	1.5000	.98404	.516	-1.3856	4.3856
Dunn-Sidak	1.00	2.00	-1.6000	1.06406		-4.0071	1.7071
		3.00	-1.0000	.98989		-4.1163	2.1163
		4.00	-2.5000	.98910		-5.0090	.0090
	2.00	1.00	1.6000	1.06406		-1.7071	4.9071
		3.00	.6000	.97163		-2.4338	3.6338
		4.00	-.9000	.98909		-3.8248	2.1248
	3.00	1.00	1.0000	.98989		-2.1163	4.1163
		2.00	-.6000	.97163		-3.6338	2.4338
		4.00	-1.6000	.98247		-4.3173	1.3173
	4.00	1.00	2.5000	.98910		-.0090	5.0090
		2.00	.9000	.98989		-2.1248	3.9248
		3.00	1.5000	.98247		-1.3173	4.3173

نتائج تحليل التباين الأحادي: اختبار Tukey واختبار Scheffe واختبار Dunnett C

للفروقات البعدية.

ان افرق بين متوسطات المجموعات 1،2،3،4 غير دال احصائياً، أي لا يوجد فروق

دات دلالة احصائية بين المجموعات الأربع.

Homogeneous Subsets

MARK

	SEC	N	Subset for alpha = .05
			1
Tukey HSD ^a	1.00	10	5.1000
	3.00	10	6.1000
	2.00	10	6.7000
	4.00	10	7.6000
	Std.		.070
Scheffe ^a	1.00	10	5.1000
	3.00	10	6.1000
	2.00	10	6.7000
	4.00	10	7.6000
	Std.		.111

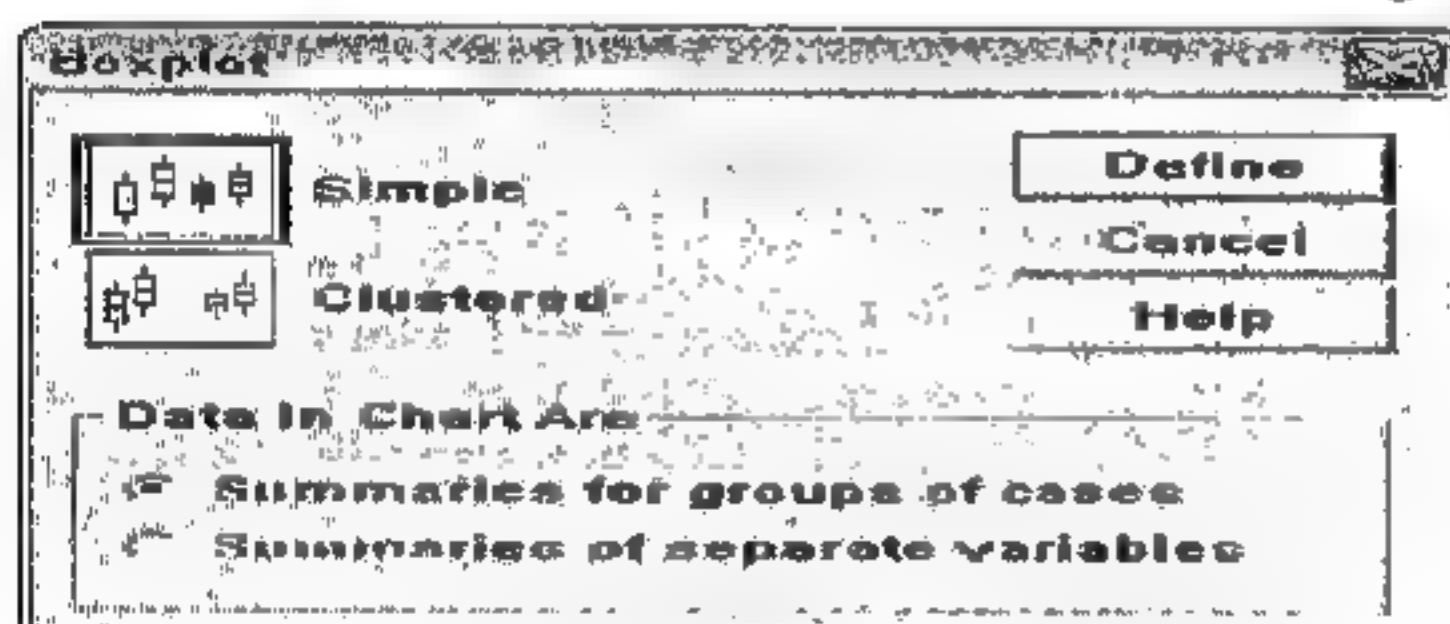
Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 10.000.

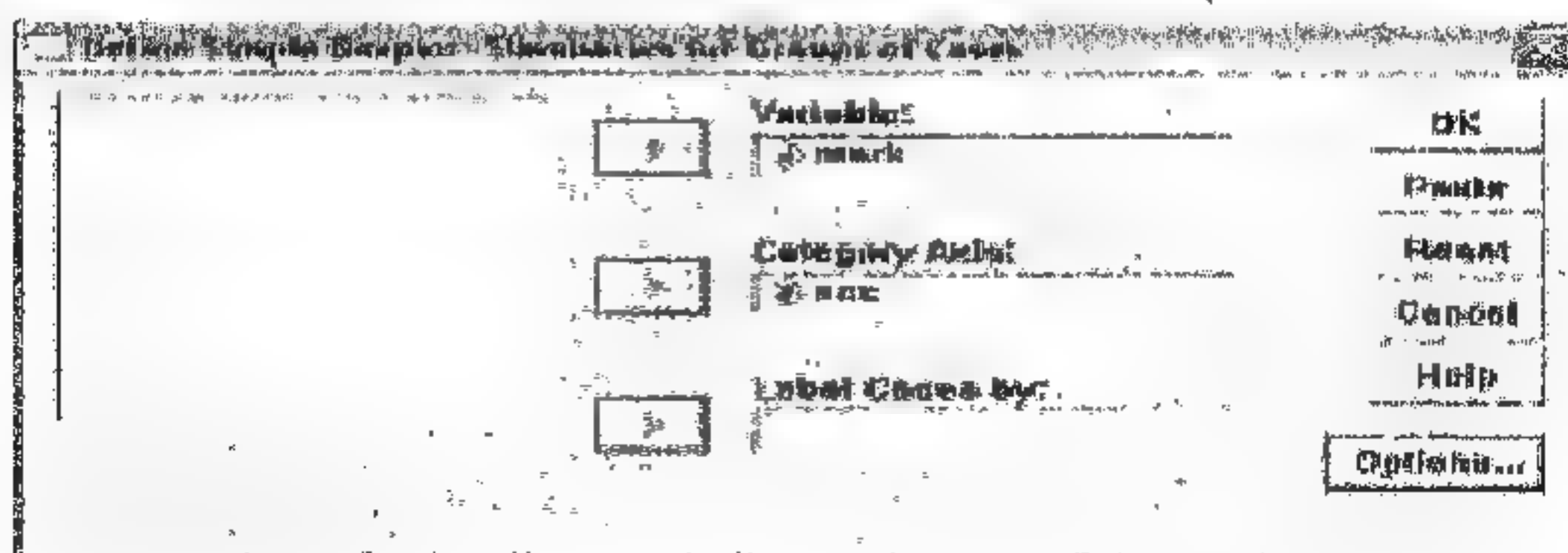
نتائج تحليل التباين الأحادي: اختبار Scheffe واختبار Tukey للفروقات البعدية للمجموعات المتماثلة.

* استخدام الرسوم البيانية لتوضيح نتائج تحليل التباين الأحادي

Graph - Boxplot... ان توزيع التغير يختلف من فئة الى اخرى



انقر Simple ثم انقر زر Define تظهر الشاشة أدناه:



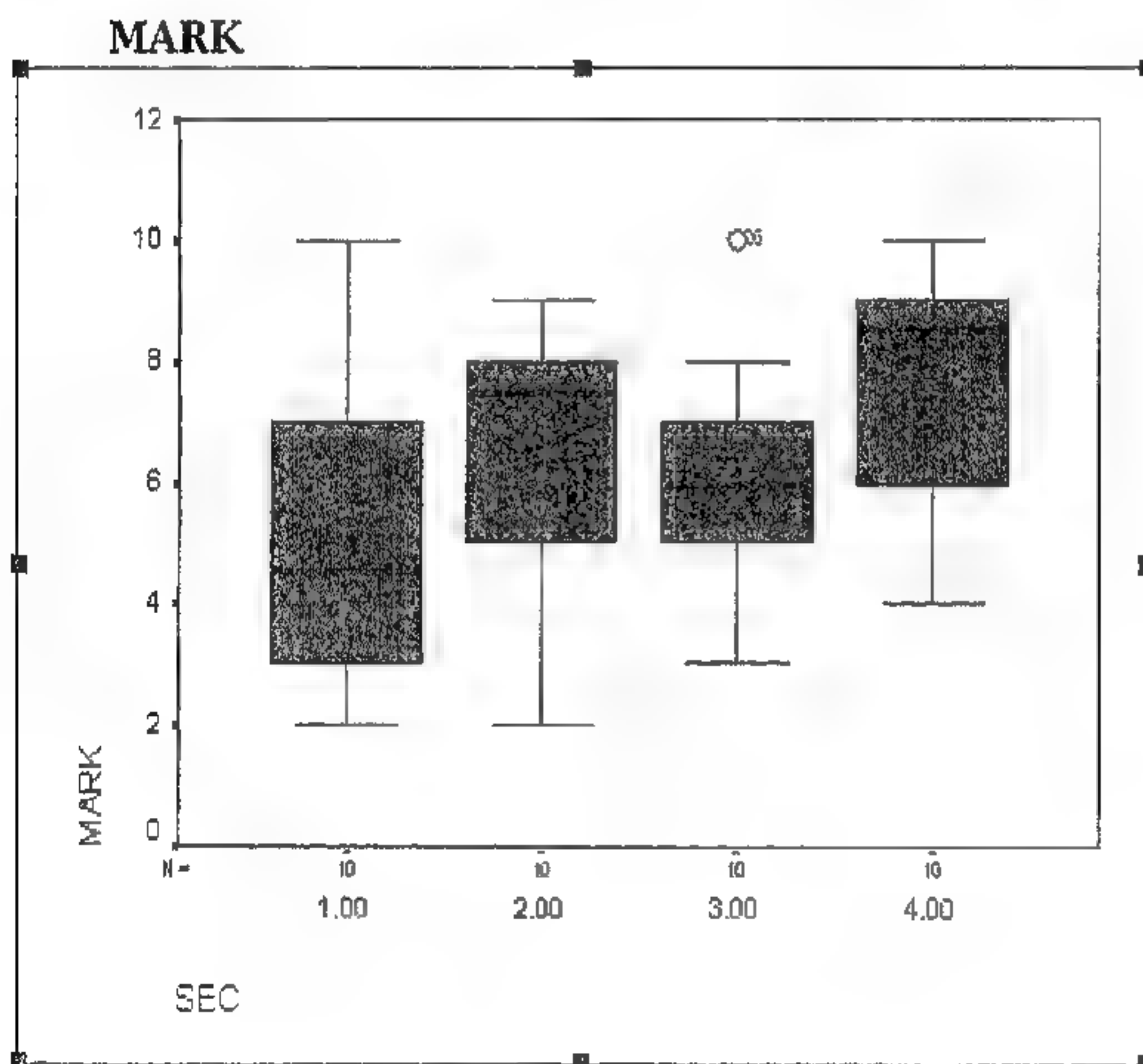
ضع المتغير mark في نافذة: Variable ، وضع المتغير sec في نافذة: Category Axis

ثم انقر زر Ok ، تظهر لك شاشة المخرجات المبينة أدناه:

SEC

Case Processing Summary

		Cases					
		Valid		Missing		Total	
SEC	MARK	N	Percent	N	Percent	N	Percent
1.00		10	100.0%	0	0%	10	100.0%
2.00		10	100.0%	0	0%	10	100.0%
3.00		10	100.0%	0	0%	10	100.0%
4.00		10	100.0%	0	0%	10	100.0%



6-9 تمارين Exercise

س1: احب عن ما يلي:

- 1- لماذا تكون قيمة F مساوية للواحد الصحيح ، عندما لا يكون هناك أثراً للمعالجة؟
- 2- وضح لماذا نستخدم اسلوب تحليل التباين الأحادي بدلاً من تكرار استخدام اختبار t لاختبار الفروق بين ثلاثة أوساط أو أكثر.

3- عسى افترض ان رئيس جامعة مؤتة بصدد ترقية دكتور لدرجة أعلى، فما هو نوع الخطأ المتوقع الوقوع فيه إذا كان:

- أ- الفرضية هي أن الدكتور يستحق الترقية، وتم قبول الفرضية الصفرية H_0 خطأً.
- ب- الفرضية هي أن الدكتور يستحق الترقية، وتم رفض الفرضية الصفرية H_0 خطأً.
- ج- الفرضية هي أن الدكتور يستحق الترقية، وتم قبول الفرضية الصفرية H_0 بشكل صحيح.
- د- الفرضية هي أن الدكتور يستحق الترقية، وتم رفض الفرضية الصفرية H_0 بشكل صحيح.

س2- اجدول التالي يمثل نتائج ثلاثة شعب مختلفة في مادة القياس والتقويم في جامعة مؤتة تم تدريسهم من قبل ثلاثة اساتذة مختلفين.

لفرض الاساتذة كالتالي: الاستاذ 1 ، 2 ، 3

الشعبة: 1 ، 2 ، 3

رقم الطالب الشعبة	الاستاذ 1 الشعبة 1	الاستاذ 2 الشعبة 2	الاستاذ 3 الشعبة 3
1	45	66	60
2	50	78	50
3	55	90	70
4	60	67	90
5	35	75	95
6	50	72	93
7	62	66	88
8	64	70	74
9	71	80	51
10	80	72	35

- 1- هل هناك فروق معنوية بين متوسطات الشعب الثلاث، اوجد جدول تحليل التباين،
- 2- اختبر الفروق بمستوى معنوية ($\alpha = 0.05$).
- 3- اختبر تجانس التباينات.
- 4- هل يوجد هناك استاذ مميز ام لا.

س3- في دراسة لمعرفة أثر المعاملات الهرمونية علي إنتاج البيض أخذت عينة لأربعة معاملات هرمونية تتمثل في 10 دجاجات لكل معاملة وتم تسجيل إنتاج البيض السوي لكل دجاجة كما هو موضح بعد:

المعاملة الأولى (X_1): 180-190-200-210-230-210-220-180-240-240 بيضة.

المعاملة الثانية (X_2): 230-260-250-270-280-290-230-240-270-280 بيضة.

المعاملة الثالثة (X_3): 130-140-150-170-190-180-200-150-210-180 بيضة.

المعاملة الرابعة (X_4): 240-250-230-210-260-220-240-260-270-220 بيضة.

والمطلوب :

- 1- وضع النموذج الإحصائي والنظرية الفرضية.
- 2- تكوين جدول تحليل التباين (ANOVA).
- 3- مقارنة متوسطات المعاملات الهرمونية الأربعة.

س4- في تجربة لاستخدام أربعة أعلاف في تغذية أربعة مجموعات تتكون كل مجموعة من خمسة كتاكيت حيث كانت الزيادة في الوزن خلال فترة ما نتجة هذه المعاملات كالتالي :

العلف	الزيادة اليومية في الوزن
1	52-21-42-49-55
2	63-98-30-112-61
3	92-95-81-97-42
4	154-85-169-137-169

والمطلوب:

1. وضع النموذج الإحصائي والنظرية الفرضية.
2. تكوين جدول تحليل التباين ANOVA
3. اختبار معنوية الفروق بين متوسطات المعاملات.

س5 : جرّبت 4 أنواع من هرمونات النمو علي أربع مجموعات من الفئران (كل مجموعة مولودة في بطن واحدة) لمدة معينة وكانت الزيادة في الوزن بعد هذه المدة هي:

هرمونات: النمو	الزيادة في الوزن بالجرام
أ	40-20-70
ب	60-40-60
جـ	50-40-80
د-	20-40-30

والمطلوب:

- 1- وضع النموذج الإحصائي والنظرية الفرضية.
- 2- تكوين جدول تحليل التباين (ANOVA) .
- 3- مقارنة متوسطات المعاملات الهرمونية الأربعة.

س6: اذا كان لديك العينات الثلاثة المبينة بالجدول ادناه:

الأولى العينة	الثانية العينة	الثالثة العينة
40	27	33
41	28	32
40.5	26.5	33.5
38.5	26.5	31.5
$X_1 = 40$ $S_1 = 1.08$	$X_2 = 27$ $S_2 = 0.71$	$X_3 = 32.5$ $S_3 = 0.91$

السؤال: هل في البيانات ما يكفي لوجود فرق بين المتوسطات؟ ولصالح أي المجموعات.

س7: في دراسة لتأثير وجود الطلاب في الصفوف على تحصيلهم في مادة الإحصاء، قام أساذ الإحصاء بأخذ عينات عشوائية ومستقلة من ثلاثة صفوف (يقوم تدريسها) كل منها مكون من خمسة طلاب وقام الأستاذ برصد درجاتهم والجدول التالي بينها. بمستوى معنوية $\alpha = 0.05$ اختبر ما إذا كان متوسط النتائج في اختبارات الأداء يختلف في تحصيل الطلاب.

الصف 1	الصف 2	الصف 3
58	96	66
62	87	65
77	66	88
90	55	92
80	78	60

- 1- هل هناك فروق معنوية بين متوسطات الصفوف الثلاث، اوجد جدول تحليل التباين. اختبر الفروق بمستوى معنوية $(\alpha = 0.05)$.
- 2- اختبر تجانس التباينات.
- 3- هل يوجد هناك صف مميز ام لا.

س8: اراد باحث ان يعرف الفعالية النسبية لأربع طرق في الدعاية والاعلان فاختر 40 زبون ووزعهم عشوائياً في اربع مجموعات (10 في كل مجموعة). كان الاعلان بالطريقة الأولى عن طريق التلفزيون، والثانية عن طريق الراديو، والثالثة عن طريق الملصقات والرابعة عن

طريق الصحف ومن ثم اخذت الاراء عن المادة موضوع الاعلان فكانت نتائجها كما يلي:

الصحف	الملصقات	الراديو	التلفزيون
66	33	55	76
88	52	76	68
63	53	59	80
61	25	86	44
53	77	53	35
51	44	42	90
45	39	51	66
46	49	73	87
70	53	45	56

افحص فرضية الباحث حول وجود اختلافات بين الطرق الاربعة في الدعاية
($\alpha = 0.05$)

س9: قسمت مدينة اربد الى اربعة مناطق وهي الحي الشرقي والغربي والشمالي والجنوبي، وتم اختيار عينة عشوائية تتكون من 10 بيوت معروضة للاجار في كل منطقة، وكانت الاجور المطلوبة لكل بيت بالدينار الاردني كما هو مبين في الجدول ادناه، والمطلوب معرفة ان كان هناك فرقاً معنوياً في معدل اجرة البيوت بين المناطق الاربعة، عند مستوى
($\alpha = 0.05$)

الحي الجنوبي	الحي الغربي	الحي الشرقي	الحي الشمالي
100	45	80	50
90	50	85	55
120	70	90	60
130	63	95	70
150	82	100	52
160	50	105	62
140	55	110	45
125	60	120	66
140	65	120	60
160	70	100	70

س10: اكمل الجدول التالي الذي يمثل خلاصة نتائج تحليل التباين لتجربة ما أجريت بهدف مقارنة أربعة مستويات معالجة علمياً بأن عدد الأفراد في كل مستوى (10).

Source of Variance	Sum of Squares (SS)	Degrees of freedom (DF)	Mean. Squares (MS)	F
مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	ف
Between Groups	SSB		MSB=SSB/C-1 15	
Within Groups	SSW 108			
Total	SST			
الكلية				

س11: في تجربة قامت بها إحدى المؤسسات الصحية لمعرفة أن كان هناك فرق في درجة الثقة بالنفس بين الأطفال المرضى والأطفال الأصحاء ، فأخذت عينة من الأطفال المرضى حجمها $n_1 = 18$ وعينة من الأطفال الأصحاء حجمها $n_2 = 18$ فكانت نتائج التجربة تشير إلى حصول النقاط التالية:

الانحراف المعياري	الوسط الحسابي	
3.9	23.5	الأطفال المرضى
3.1	27.8	الأطفال الأصحاء

والمطلوب إجراء الاختبار عند مستوى معنوية $\alpha = 0.01$.

س12) لمعرفة الثقل المحوري للشاحنات المارة على طريقين تم إنشاؤهما حديثاً لأحدى البلديات والمصممة لنفس المواصفات أخذت عينة ستكون من 31 شاحنة من كل طريق واتضح بأن متوسط الحمولة لها والانحراف المعياري هي كالآتي:

الانحراف المعياري	الوسط الحسابي	
3.9	23.5	الطريق الأول
3.1	27.8	الطريق الثاني

فهل نستدل على وجود فروقات في حمولة الشاحنات المارة على كلا الطريقتين عند مستوى معنوية $\alpha = 0.10$

س13: معمل فيه خطين إنتاجيين لإنتاج نوعين من المصابيح الكهربائية أخذت عينة من الخط الأول حجمها (60) مصباح فكانت نسبة المصابيح الغير صالحة 12% بينما في العينة الثانية التي أخذت من الخط الثاني والتي كان حجمها (80) مصباح كانت نسبة المصابيح الغير صالحة 9%.

فإذا كان متوسط عمر المصابيح في عينة الأولى (900) ساعة والانحراف المعياري (20) وفي العينة الثانية كان المتوسط (970) ساعة والانحراف المعياري (17) . والمطلوب إيجاد:
أ- تقدير الفترة للفرق بين متوسطي عمر المصابيح المنتجة في الخطين بمعامل ثقة (90%).

ب- تقدير الفترة للفرق بين نسبي المصابيح الصالحة في خطين الإنتاجين .

س14: اختبرت عينة تتكون من 50 مدرسة إعدادية ، فكان معدل الدرجة النهائية للطلاب لهذه المدارس هو 61 وبانحراف معياري مقداره 4.5 درجة . في حين أوضحت أمانة التربية المسؤولة عن هذه المدارس بأن المعدل النهائي يزيد على 62 وبانحراف معياري مقداره 5.1. والمطلوب اختبار مدى صحة ادعاء أمانة التربية عند مستوى معنوية $\alpha = 0.05$

س15: اختبرت عينتين من طلبة إحدى الجامعات ، الأولى تمثل الذكور والثانية إناث ، وطبقت عليهم إحدى أساليب التدريس الحديثة. وفي نهاية السن الدراسية كان عدد الناجحين لكل من العينتين هو كما يلي:

النوع	حجم العينة	عدد الناجحين
ذكور	86	50
إناث	60	45

فهل نستدل بأن هناك فرق بين الذكور والإناث في النتيجة النهائية عند مستوى معنوية

$$\alpha = 0.01$$

س16: في دراسة قامت بها إدارة التلفزيون لمعرفة إن كان برنامجها الترفيهي له نفس الأهمية لكافة الفئات العمرية ، فأخبرت عينة من المشاهدين ، وحصلت على النتائج المبينة في الجدول التالي . المطلوب اختبار أن كان هناك فروق بين الفئات العمرية اتجاه البرنامج المعينة عند مستوى معنوية $\alpha = 0.10$.

المجموع	60 سنة فأكثر	40-59	18-39	أقل من 18	الرغبة
78	11	17	18	32	لا يرغب
307	36	71	85	115	يرغب
153	28	22	41	62	يرغب جدا
538	75	110	144	209	المجموع

س17: للمقارنة بين عدد الكيلو مترات التي يمكن قطعها بالغالون الواحد لنوعين من البنزين، تم اختيار سيارتي صالون من نفس النوع والموديل ، وتم تسير كل منها عدة مرات فكانت النتيجة كالآتي:

عدد الكيلو مترات المقطوعة	
لنوع الأول من البنزين	لنوع الثاني من البنزين
78	70
80	81
76	79
81	72
85	69
78	75
82	66
83	70
86	68
90	75

والمطلوب اختبار ان كان هناك فرق بين تباين المسافات المقطوعة لكل من نوعي البنزين المستخدم عند مستوى معنوية $\alpha = 0.05$.

س18: المعطيات الآتية تمثل طول شعرة القطن (ملم) لخمسة أصناف من القطن لعينة عشوائية حجمها 10 شعيرات من كل صنف . المطلوب اختبار الفرض القائل بأنه لا توجد فروق جوهرية بين هذه الأصناف الخمسة تحت مستوى معنوية $\alpha = 0.01$.

الصنف	طول شعرة القطن (ملم)										مجموع
A	39	39	38	37	36	37	39	39	38	39	381
B	39	40	38	41	38	39	41	41	38	38	392
C	35	36	36	37	35	36	35	35	38	37	360
D	36	40	38	39	39	41	42	41	40	40	396
E	35	36	35	34	33	32	33	32	32	34	336

س19: اختبر فرضية $H_0: \mu_1 = \mu_2 = \mu_3$ للمعطيات في الجدول الآتي :

الصنف	المشاهدات						المجموع
1	4	7	6	3			20
2	7	8	6	6	5	4	36
3	5	6	7				18

س20: استخدمت أربع طرق لأربع مجاميع من الطلبة لتعليمهم جدول الضرب، وكانت النتائج كما يلي ، المطلوب اختبار فيما إذا كانت فروق جوهرية بين الطرق الأربع وعند المستوى معنوية $\alpha = 0.05$.

طريقة التعليم	المشاهدات						المجموع
1	7	6	8	5	9	7	42
2	8	9	10	7	8	6	48
3	7	8	10	5	6	3	39
4	8	6	5	4	9	4	36

س21: المعطيات في الجدول الآتي تمثل عاملين هما ويضم أربعة أنواع من الأمراض ويضم أربعة من المستشفيات والقيم تمثل عدد الأيام الذي يستغرقه المريض للعلاج في المرض. المطلوب اختيار الفرضيات الآتية:

أ- هل هناك اختلاف في مدة العلاج باختلاف المستشفى

ب- هل ان نوع المرض له علاقة بفترة العلاج ؟

ج- هل هناك تفاعل بين نوع المستشفى ونوع المرض ؟

المجموع	4	3	2	1	العامل B
					العامل A
534	82	24	25	20	1
	13	28	30	25	
	62	24	29	22	
	92	25	28	27	
	23	30	20	21	
	146	131	142	115	
765	40	39	30	30	2
	45	42	29	54	
	50	36	31	30	
	45	42	30	53	
	60	40	30	63	
	24	199	150	176	
766	42	41	32	31	3
	50	45	35	30	
	40	40	30	40	
	55	40	34	35	
	45	35	30	30	
	232	201	137	166	
509	29	24	33	20	4
	30	25	25	21	
	28	30	28	20	
	27	26	30	20	
	30	23	31	19	
	144	128	137	100	
2574	762	659	596	557	المجموع الكلي

س22: أجريت تجربة لبيان تأثير أربعة أنواع في الاغذية في زيادة وزن مجموعة من الابقار تنتمي لثلاث سلالات مختلفة وتم اعطاء كل نوع في الابقار الاغذية الى خمسة ابقار في كل سلالة وكانت النتائج التالية التي تمثل مجموع الزيادة في وزن الابقار الخمسة لكل سلالة ولكل نوع من الغذاء والمطلوب:

- 1- تكوين جدول تحليل التباين
- 2- إجراء كافة الاختبارات الممكنة عند مستوى معنوية 0.05
- 3- في حالة وجود تأثير معنوية لنوع الغذاء على زيادة الوزن حدد أي الأنواع كانت السبب.

السلالة	نوع الغذاء				المجموع
	A	B	C	D	
1	91	98	112	109	408
2	113	116	114	119	162
3	127	121	116	121	185
المجموع	331	333	342	349	755

س23) في تجربة طبية معينة اخذت مجاميع من المرضى واعطيت جرعات بمستويات مختلفة من الدواء معين وبعد فترة زمنية معينة تم حساب نسب الشفاء للمجاميع وكانت كما يلي:

المستوى	نسب الشفاء						المجموع
A	81	80	75	80	70	70	456
B	70	70	68	65	69	70	412
C	62	63	64	67			256
D	68	66	65	63			262

- 1- اوجد جدول تحليل التباين للبيانات اعلاه واختبر فيما اذا كانت هناك فروق معنوية بين مستويات الجرعة عند مستوى معنوية (5%)
- 2- في حالة وجود فروق معنوية حدد ايا من مستويات الجرعة هي السبب في الفرق معنوي وعلق على النتيجة .

24: تمرين (17.9): اكمل الجدول تحليل تباين التالي واكتب النموذج الملائم ثم اختبر الفرضية القائلة بوجود تأثير مشترك للصفوف والاعمدة بمستوي معنوية $\alpha = 0.05$

مصدر تباين	درجة الحرية	مجموع المربعات	متوسطات المربعات	F
بين الاعمدة		12.25		1.306
بين الصفوف			167.25	53.52
تفاعل	9			
الخطأ				
الكلي	31	598		

الفصل الخامس عشر

المقارنات المتعددة

Multiple Comparisons

1-10 مقدمة

2-10 أنواع المقارنات المتعددة

1- المقارنات المخطط لها

أ- طريقة المقارنات المتعامدة Orthogonal

ب- طريقة دن Dunn وتسمى أيضاً طريقة بنفوروني

Bonferroni

2- المقارنات غير المخطط لها

أ- طريقة شافيه Scheffe

ب- طريقة توكي Tukey

ج- طريقة نيومان كولز Newman Kuelz

3-10 تمارين Exercise

الفصل العاشر

المقارنات المتعددة

Multiple Comparisons

1-10 مقدمة

تحليل التباين **Analysis of Variance**: تعميم لاختبار t ويستخدم لاختبار ثلاث عييات أو أكثر، ويستخدم لمقارنة عدة متوسطات حسابية في آن واحد $H_0: \mu_1 = \mu_2 = \dots = \mu_k$ ، وهو طريقة ذكية لاختبار اختلاف أوساط أكثر من مجموعتين دفعة واحدة من خلال التباين.

المقارنات البعدية **Post hoc Comparisons**: وهي عبارة عن اختبارات تتبعية تحتاج إليها عندما يوجد أكثر من وسطين وهذه المتوسطات مختلفة من الناحية الإحصائية (أي رفضت الفرضية الصفرية).

نلجأ للمقارنات البعدية عند رفض فرضية صفرية تتضمن ثلاثة متوسطات أو أكثر.

2-10 أنواع المقارنات المتعددة

يمكن تصنيف المقارنات المتعددة إلى نوعين رئيسيين هما:

- 1- المخطط له أو القبلي **Planned or Apporiori** ويتم الترتيب له بشكل مسبق قبل إجراء اختبار F وإن استخدامه لا يتطلب أن تكون F ذات دلالة إحصائية.
- 2- غير المخطط له أو البعدي **Post hoc or Apostriori** يتطلب أن تكون F ذات دلالة إحصائية قبل إجراءه.

1- المقارنات المخطط لها

وتستخدم بدون إجراء تحليل التباين للبيانات وهي تعتمد على إجراء اختبار T -test ومن هذه الطرق ما يلي:

أ- طريقة المقارنات المتعامدة Orthogonal

من قائمة Analyze - Descriptive Statistics - Crosstabs

Case Processing Summary						
		Cases				
		Valid		Missing		Total
		N	Percent	N	Percent	
MARK * SEC		40	100.0%	0	0%	40

MARK * SEC Crosstabulation						
		SEC				Total
		1.00	2.00	3.00	4.00	
MARK	2.00	Count 1	1			2
		% within MARK 50.0%	50.0%			100.0%
	3.00	Count 2		1		3
		% within MARK 66.7%		33.3%		100.0%
	4.00	Count 2	1	1	1	5
		% within MARK 40.0%	20.0%	20.0%	20.0%	100.0%
	5.00	Count 1	1	2	1	5
		% within MARK 20.0%	20.0%	40.0%	20.0%	100.0%
	6.00	Count 1		2	1	4
		% within MARK 25.0%		50.0%	25.0%	100.0%
	7.00	Count 2	2	2	1	7
		% within MARK 28.6%	28.6%	28.6%	14.3%	100.0%
	8.00	Count 3		1	1	5
		% within MARK 60.0%		20.0%	20.0%	100.0%
	9.00	Count 2			4	6
		% within MARK 33.3%			66.7%	100.0%
	10.00	Count 1		1	1	3
		% within MARK 33.3%		33.3%	33.3%	100.0%
Total	Count	10	10	10	10	40
	% within MARK	25.0%	25.0%	25.0%	25.0%	100.0%

Chi-Square Tests			
	Value	df	Asymp Sig (2-sided)
Pearson Chi-Square	21.429 ^a	24	.613
Likelihood Ratio	28.693	24	.318
Linear-by-Linear Association	4.477	1	.034
N of Valid Cases	40		

a. 38 cells (100.0%) have expected count less than 5. The minimum expected count is .50.

بما أن قيمة مربع كاي = 21.429 وهذا يعني أن هناك علاقة قوية موجبة ولكن ليست

ذات دلالة احصائية حيث أن مستوى المعنوية = 0.613

ب- طريقة دن Dunn وتسمى أيضاً طريقة بنفروني Bonferroni

Oneway

ANOVA

MARK

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	33.075	3	11.025	2.277	.096
Within Groups	174.300	36	4.842		
Total	207.375	39			

Post Hoc Tests

Multiple Comparisons

Dependent Variable: MARK

Bonferroni

J SEC	K SEC	Mean Difference (J-K)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1.00	2.00	-1.0000	.98404	.676	-4.3474	1.3474
	3.00	-1.0000	.98404	1.000	-3.7474	1.7474
	4.00	-2.5000	.98404	.093	-5.2474	.2474
2.00	1.00	1.0000	.98404	.676	-1.1474	4.3474
	3.00	.0000	.98404	1.000	-2.1474	3.3474
	4.00	-.9000	.98404	1.000	-3.6474	1.8474
3.00	1.00	1.0000	.98404	1.000	-1.7474	3.7474
	2.00	-.6000	.98404	1.000	-3.3474	2.1474
	4.00	-1.5000	.98404	.917	-4.2474	1.2474
4.00	1.00	2.5000	.98404	.093	-.2474	5.2474
	2.00	.9000	.98404	1.000	-1.8474	3.6474
	3.00	1.5000	.98404	.917	-1.2474	4.2474

2- المقارنات غير المخطط لها

نعمل في البداية تحليل تباين، ثم نعمل مقارنات بعدية حسب القرار.

* الحصول على الفروق الحرجة:

هناك أكثر من آلية للحصول على الفروق الحرجة ولكننا سنهتم بثلاثة طرق وهي:

أ- طريقة شافيه Scheffe

ب- طريقة توكي Tukey

ج- طريقة نيومان كولز Newman Kuelz

أ- طريقة شافيه Scheffe

تستخدم هذه الطريقة في إجراء جميع المقارنات بين الأوساط وهي مفضلة على الطرق الأخرى عندما تكون أحجام الخلايا غير متساوية وفي إجراء المقارنات المعقدة، وحتى تكون

المقارنة ذات دلالة احصائية فإنها يجب أن تكون أكبر من المقدار CDS

الفرق الحرج عند شافيه هو CDS (Critical Def. Scheffe)

$$CDS = \sqrt{(k-1) \alpha F_{k-1, n-k} MSE (1/n_i + 1/n_j)}$$

F: القيمة الحرجة عند مستوى الدلالة المعني

K: عدد المجموعات

MSE: Error with in main sequare

الاختلاف بين موقع وآخر فيما يتعلق بالفروق الحرجة يكمن فقط في اختلاف أحجام

العينات، وفي حالة تساوي الأحجام للمجموعات المختلفة يكفي فرق حرج واحد.

$$CDS = \sqrt{2(3.29) (1.83)(1/12 + 1/12)} = 1.42$$

أي فرق يزيد عن هذه القيمة الحرجة يكون دال احصائياً.

هناك فرق بين X'_1 و X'_3 دال احصائياً وهو لصالح X'_1 لأنها أكبر.

هناك فرق بين X'_2 و X'_3 دال احصائياً وهو لصالح X'_2 لأنها أكبر.

الفرق بين X'_1 و X'_2 غير دال احصائياً.

طريقة شافيه غير حساسة للفروق (متشددة)، بمعنى لا تعتمد الفرق أن يكون فرق إلا إذا كان ذا دلالة قوية، ولذلك نتوقع أن يكون الفرق الحرج عند شافيه أكثر منه في الطرق الأخرى.

ب- طريقة توكي Tukey أو المسماة Honestly Significant Difference (HSD)

يعطى الفرق الحرج في طريقة توكي Tukey حسب القانون التالي:

$$CDT = \alpha q_{k,v} \sqrt{MSE (1/2n_i + 1/2n_j)}$$

k: عدد المجموعات.

v: درجات الحرية للمقام.

$$CDT = 0.05 q_{3,33} \sqrt{1.83(1/24 + 1/24)} = 3.49 \sqrt{1.83(2/24)} = 1.36$$

أي فرق أكثر من هذه القيمة الحرجة يكون دال احصائياً.

جميع الفروق دالة احصائياً.

ج- طريقة نيومان كولز Newman-Kuelz

وتفيد هذه الطريقة في المقارنات بين أزواج الأوساط فقط.

يعطى الفرق الحرج في طريقة نيومان كولز Newman Kuelz حسب القانون التالي:

$$CDNs = \alpha q_{c,v} \sqrt{MSE (1/2n_i + 1/2n_j)}$$

c: عدد المتوسطات ضمن مدى موضع الاهتمام (المقارنة)

v: درجات الحرية للمقام.

$$X'_1, X'_2, X'_3 \text{ تكون } c=3, X'_1 - X'_2, X'_1 - X'_3, X'_2 - X'_3$$

ملاحظة: في حالة نيومان كولز نحتاج الى أكثر من فرق حرج حتى في حالة تساوي

الحجوم.

$$\text{بين } X'_2 - X'_3 \text{ تكون } c=3$$

نفس قيمة توكي $CDNs = CDT = 1.36$ لأن قيمة $k = c$

$$\text{بين } X'_1 - X'_3 \text{ تكون } c = 2 \text{ وبين } X'_1 - X'_2 \text{ تكون } c = 2$$

$$CDNs = 0.05 q_{2,33} \sqrt{1.83 (1/24 + 1/24)} = 1.13$$

جميع الفروق دالة احصائياً.

مثال: في دراسة تشتمل على 5 مجموعات كانت النتائج كما في الجدول التالي:

المجموعة	\bar{X}	S^2	N
G1	50	48	20
G2	52	50	25
G3	55	51	25
G4	49	51	25
G5	44	50	25

1- هل توجد فروق دالة إحصائية بين متوسطات المجموعات الخمس عند مستوى

الدلالة الإحصائية $\alpha = 0.05$ ؟

2- هل يلزم مقارنات بعدية ؟

3- إذا لزم استخدم طريقة توكي وطريقة شافيه وطريقة نيومان كورلر لتحديد فيما إذا

كان الفرق بين متوسطات المجموعة الأولى والمجموعة الخامسة دال إحصائياً عند

مستوى الدلالة الإحصائية $\alpha = 0.05$ ؟

4- هل توجد فروق دالة إحصائية بين متوسطات المجموعات الخمس عند مستوى

الدلالة الإحصائية $\alpha = 0.05$ ؟

المجموعة	\bar{X}	S^2	N	$\bar{X} * n$	SSB	SSW
G1	50	48	20	1000	0	912
G2	52	50	25	1300	100	1200
G3	55	51	25	1375	625	1224
G4	49	51	25	1225	25	1224
G5	44	50	25	1100	900	1200
المجموع			120	6000	1650	5760
\bar{X}_w	50	SSB	1650		SSW	5760

1- نجد الوسط الموزون

$$\bar{X} = \frac{(20*50)+(25*52)+(25*55)+(25*49)+(25*44)}{120} = 50$$

$$SSB = 20(50-50)^2 + 25(52-50)^2 + 25(55-50)^2 + 25(49-50)^2 + 25(44-50)^2 = 1650$$

$$SSW = 19(48) + 24(50) + 24(51) + 24(51) + 24(50) = 5760$$

Source of Variance	Sum of Squares (SS)	Degrees of freedom (DF)	Mean. Squares (MS)	F
مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	ف
Between Groups بين المجموعات	1650	5 - 1 4	MSB=SSB/C-1 =1650/4 = 412.5	MSB/MS W 412.5/50.09 8.24
Within Groups ضمن المجموعات	SSW 5760	N - C =120-5 = 119 115	MSW=SSW/N-C =5760/120-5 = 50.09	
Total الكل	SST 7410	N - 1 =120-1 =119 119		

يُحد قيمة F الحرجة $F_{0.05,4,115} = 2.45$

بما أن قيمة F المحسوبة 8.24 أكبر من قيمة F الحرجة $F_{0.05,4,115} = 2.45$ لذلك

نرفض الفرضية الصفرية. أي أن هناك فروق بين متوسطات المجموعات.

2- هل يلزم مقارنات بعدية ؟

بما أن هناك فروق بين متوسطات المجموعات إذا يلزم مقارنات بعدية.

3- إذا لزم استخدم طريقة توكي وطريقة شافيه وطريقة بيومان كولر لتحديد فيما إذا

كان الفرق بين متوسطات المجموعة الأولى والمجموعة الخامسة دال احصائياً عند مستوى الدلالة

الاحصائية $\alpha = 0.05$ ؟

جدول الفروق بين المتوسطات

	$X'_5 = 44$	$X'_4 = 49$	$X'_1 = 50$	$X'_2 = 52$	$X'_3 = 55$
$X'_5 = 44$	-	5	6	8	11
$X'_4 = 49$		-	1	3	6
$X'_1 = 50$			-	2	5
$X'_2 = 52$				-	3
$X'_3 = 55$					-

1- طريقة شافيه Scheffe

$$CDS = \sqrt{(k-1) \alpha F_{k-1, n-k} MSE (1/n_i + 1/n_j)}$$

$$CDS = \sqrt{(5-1) (2.45) (50.09) (1/20 + 1/25)} = 6.25$$

بما أن 6 أقل من 6.65 إذا الفرق بين $X'1 - X'5$ غير دال احصائياً.

2- طريقة توكي Tukey

$$CDT = \alpha q_{k,v} \sqrt{MSE (1/2n_i + 1/2n_j)}$$

$$CDT = 0.05 q_{5,115} \sqrt{50.09 (1/40 + 1/50)}$$

$$CDT = 3.92 \sqrt{50.09 (1/40 + 1/50)} = 5.92$$

بما أن 6 أكبر من 5.92 إذا الفرق بين $X'1 - X'5$ دال احصائياً.

3- طريقة نيومان كولز Newman Kuelz

$$CDNs = \alpha q_{c,v} \sqrt{MSE (1/2n_i + 1/2n_j)}$$

$$CDNs = 0.05 q_{3,115} \sqrt{50.09 (1/40 + 1/50)} = 3.36 (1.51) = 5.07$$

بما أن 6 أكبر من 5.07 إذا الفرق بين $X'1 - X'5$ دال احصائياً.

3-10 تمارين Exercise

س1: متى نلجأ الى المقارنات البعدية؟

س2: اذكر انواع المقارنات البعدية.

س3: اشرح المقارنات المخطط لها

طريقة المقارنات المتعامدة Orthogonal

طريقة دن Dunn وتسمى ايضاً طريقة بنفوروني Bonferroni

س4: اشرح المقارنات غير المخطط لها

طريقة شافيه Scheffe

طريقة توكي Tukey

طريقة نيومان كولز Newman Kuelz

الفصل الحادي عشر

التحليل العاملي

Factor Analysis

- 1-11 مقدمة
- 2-11 مفهوم التحليل العاملي
- 3-11 أهمية التحليل العاملي وميادينه
- 4-11 أهداف التحليل العاملي
- 5-11 خطوات استخدام التحليل العاملي
- 6-11 تنفيذ التحليل العاملي من خلال برنامج SPSS
- 7-11 استخدام برمجية SPSS في حل المسائل.
- 8-11 تمارين Exercise

الفصل الحادي عشر

التحليل العاملي

Factor Analysis

1-11 مقدمة Introduction

يعتمد التحليل الإحصائي بصفة عامة على نوع المشكلة محل القياس وخصائصها، ونوع البيانات سواء كانت إسمية أو ترتيبية، أو فترية أو نسبية، وأيضاً على الأهداف المراد تحقيقها من الدراسة، وعلى ذلك فالتحليل الذي يستخدم لمشكلة معينة، قد لا يستخدم لمشكلة أخرى. أصبح التحليل العاملي يحتل مكانة هامة في البحوث بمختلف أنواعها، حيث إن هذه العلوم تخضع لكثير من المتغيرات المتداخلة، التي يكون بينها مجموعة من الارتباطات السلبية أو الإيجابية.

والتحليل العاملي أسلوب إحصائي يساعد الباحث في دراسة المتغيرات المختلفة بقصد إرجاعها إلى أهم العوامل التي أثرت فيها، فمن المعروف أن أي ظاهرة من الظواهر تنتج من عدة عوامل كثيرة، وتعتبر الظاهرة محصلة لهذه العوامل جميعاً (باهي وعبد الفتاح، 2006: 187).

أيضاً فإن التحليل العاملي أسلوب إحصائي يعمل على تجميع متغيرات ذات طبيعة واحدة في تركيبة متجانسة مرتبطة داخلياً فيما بينها في تكوين يسمى عامل بحيث يرتبط كل متغير من هذه المتغيرات بهذا العامل، أي أن كل متغير من هذه المتغيرات يتشعب على هذا العامل بقيم متفاوتة توضح الأهمية النسبية لكل متغير من هذه المتغيرات المرتبطة بالنسبة لهذا العامل (إبراهيم، 2002: 196).

ويشير عبد الخالق (1987) أن هناك اتجاه آخر عكس هذا الرأي وهو أن التحليل العاملي يقترح فروضاً، وكلما نجح في القيام بهذه المهمة انتهت وظيفة الوصف ليصبح جزءاً من النظرية السيكلوجية من حيث هو إحصاء يختصر العلاقات بين مجموعة من المتغيرات ويقترح علاقات سببية لم يسبق اكتشافها.

وان توليد الفروض ليس حكراً على التحليل العاملي فهو يشبه في ذلك طريقة الملاحظة والعمل الاكلينيكي، الا ان الاخيرين يقلون عنه في درجة الدقة والصرامة وقد يسهل تكوين الفروض في مجال تتوفر فيه ملاحظات كثيرة، الا ان اسهامات التحليل العاملي تصبح مهمة جداً في المجالات الجديدة نسبياً، وذلك في الاسراع في تكوين فروض معقولة وقيمة واستبعاد الفروض الضعيفة، ويتصل هذا الهدف باثبات الفروض او رفضها وبخاصة الفروض المتعلقة بتركيب الشخصية وتنظيمها كفروض الانماط والسمات ، مما يصعب احباطه او دحضه بالطرق غير العاملية. ويين مستوى اقتراح الفروض والتحقق منها متينة .

ومن خلال ما سبق نجد ان العوامل تكون :

- مفاهيم احصائية بحتة.

- مبادئ للتصنيف.

- وسيلة لظهور العلاقات السببية (عبد الخالق ، 1987 : 100-101) .

ويذكر ريتشارد وآخرون (1992) ان الغرض الجوهرى من التحليل العاملي هو بقدر الامكان وصف علاقات التباین التلازمی بين العديد من المتغيرات بدلالة قليلة نسبياً مع متغيرات اخرى في مجموعات مختلفة، ويمكن تصور ان هذه المجموعة من المتغيرات تمثل بنية اساسية متفردة تسمى عامل. (Johnson, Richard A. & Wichern , Dean .1992 : 296-297)

ويذكر جيلفورد (1961) انه اذا كان التحليل العاملي افضل اداة فعالة لاستخراج المعلومات من البيانات فانه ينبغي ان نكون على علم بان التحليل العاملي ليس له قوة سحرية تكشف عن تلك المعلومات التي تتضمنها البيانات المتجمعة ودائماً ينبغي على كل من يحاول استخدام الكشف عن معلومات سيكولوجية تتعلق بالتنظيم العقلي المعرفى أو سمات الشخصية ان يبدأ بفروض واضحة قابلة للاختبار (Guilford, J.P., 1961 : 231)

ونظراً لتعدد أهداف البحث وتعدد متغيراته الاجتماعية والاقتصادية، فإنه تم اختيار الاساليب الإحصائية التي تتفق مع طبيعة البيانات وتسمح بتحقيق الهدف الأول من البحث وهو تحديد الهيكل البنائي الاجتماعي والاقتصادي هي أساليب التحليل متعدد المتغيرات وهي أكثر الأساليب ملاءمة للمتغيرات والعلاقات المحددة لتكوين هذا الهيكل، أما أساليب التحليل التي تحقق الهدف الثاني من البحث وهي إعادة توزيع عبء نظام التأمين الاجتماعي فهي النماذج

الرياضية المستخدمة في بحوث العمليات والتي تركز على إعادة توزيع الموارد وهي أكثر ملاءمة لتحقيق التوزيع الأمثل لهذا العبء.

التحليل العاملي هو أسلوب إحصائي متعدد المتغيرات، يسعى إلى تحديد الأبعاد أو العوامل التي تساعد في وصف ظاهرة معقدة، عن طريق تحليل مصفوفة الارتباط (معاملات الارتباط البسيطة) بين المتغيرات المختلفة الداخلة في وصف الظاهرة، وصولاً إلى عوامل Factors محددة تكمن وراء طبيعة العلاقات الداخلية بين مجموعة المتغيرات في هذه الدراسة. وفي سبيل ذلك يسعى التحليل العاملي إلى تقليل البيانات Data Reduction بتحديد عدد العوامل القليلة التي تفسر معظم التباين في عدد كبير من المتغيرات، فبدلاً من أن يكون لدينا (35) متغيراً مثلاً يختصر التحليل العاملي هذا العدد إلى (06) عوامل مثلاً. وعادة ما تكون البيانات هي قيم (درجات أفراد على متغيرات نفسية أو اجتماعية أو تربوية).

يمثل أسلوب التحليل العاملي خطأً (مخطأً) من البحث مختلف تماماً عن الطرق الرياضية الأخرى في العلوم الاجتماعية، لاعتماده على افتراضات إحصائية، وهي نظرية شائعة ومفضضة لدى كثيرين، لأنها تحاول أن تجيب على السؤال الذي طالما سألته العلم: ما هو أقل عدد من المفاهيم التي يمكن أن تنظم تعقد الظاهرة وتصفها؟ هيم (عبدالحالوق، 1994).

يظهر من هذا أننا نستطيع أن نستخدم هذا الأسلوب الإحصائي في تنظيم مجال جديد يحتاج للتعرف على خصائصه ومتغيراته، وهي حاجة يسعى إليها الباحث عندما يطرق مجالاً جديداً لا يعرف كل متغيراته أو مدى تعلق المتغيرات المختلفة بظواهره الرئيسية، والنتيجة المباشرة لهذه الخطوة الاستكشافية هي إعادة الدراسة والتناول للمتغيرات المهمة في المجال، وبناء الفروض التي تفسر العلاقات بين هذه المتغيرات. وتستند فلسفة التحليل العاملي إلى تحليل الارتباطات بين المتغيرات، بغرض استخلاص أقل عدد ممكن من العوامل التي تعبر عن أكبر قدر من التباين بين المتغيرات. وبذلك يبدأ التحليل العاملي بحساب معاملات الارتباط البسيطة بين المتغيرات وتسجيلها في مصفوفة تصلح لهذا النوع من التحليل، ويكون الهدف هو توضيح وتفسير العلاقات بين تلك المتغيرات، وينتج عنها عدد قليل من المتغيرات الجديدة (المفترضة) تسمى بالعوامل التي تحتوي على كل المعلومات الأساسية. وينتهي التحليل بمصفوفة عوامل ما قبل التدوير ومصفوفة عوامل بعد التدوير، تلك العوامل التي أدت إلى ذلك الارتباط.

ويمكن القول بأن التحليل العاملي نشأ في كنف علم النفس، حيث البدايات الأولى على يد الرواد الأوائل لعلم النفس من أمثال ثورنديك Thorndike وبيرسون Pearson وهوتنج Hotteling وطومسون Tomson وجيلفورد Gulford وجالتون Galton وهولزنجر K.J. Holzinger وبييرت C. Burt وثيرستون L.L. Thurston وإلكسندر... W.P. Alexander إلخ. ثم انتقل إلى التطبيقات العملية والعلمية في شتى فروع المعرفة (باهي، 2002).

ويرجع الفضل في ذلك إلى سبيرمان C. Spearman منذ عام 1863م، بأن طور أفكاره وأضاف أبعاداً جديدة للمفهوم في دراسته التي نشرها عام 1904م، حيث بين أن العامل هو السبب في الارتباط الموجب بين أي ظاهرتين، وفي تطور لاحق أعلن سبيرمان أن العامل هو السبب المباشر لوجود الارتباطات الموجبة القائمة بين أي عدد من المتغيرات أو المقاييس. وفرق سبيرمان بين عاملين هما: العامل العام General Factor، وهو العامل المشترك بين جميع المتغيرات، والعامل الخاص Specific Factor، وهو الذي يميز النواحي الخاصة التي ينفرد بها المتغير عن غيره من المتغيرات الأخرى. ولذا فمعامل ارتباط أي عاملين خاصين يساوي الصفر. ولذلك سميت نظرية سبيرمان العاملية بنظرية العاملين، وقد عدل بعض العلماء، مثل هولنجر، بنظرية العاملين فأضاف لها نوعاً من العوامل التي توجد في طائفة من المتغيرات دون غيرها، وسماها بالعوامل الطائفية Group Factor وطبقاً لهذه النظريات نجد أنه من الممكن تصنيف العوامل التي يتوصل إليها الباحثون في البحوث النفسية والتربوية والاجتماعية إلى ثلاثة أنواع هي:

1- العامل العام: هو العامل المشترك الذي يوجد في جميع المتغيرات (الاختبارات) التي تخضع لتحليل العاملي، ويعبر عنه في هذه الحالة بالنمط العام، كما هو الحال في الذكاء العام على سبيل المثال.

2- العامل الطائفي: هو العامل الذي يوجد في بعض المتغيرات (الاختبارات) التي تخضع للتحليل، وليس فيها كلها، وهو يفسر ارتفاع قيم معاملات الارتباط بين الاختبارات التي تقيس الذكاء، ومن أمثلة العوامل الطائفية القدرات العقلية الموجودة في الذكاء على سبيل المثال، مثل القدرة المكانية أو القدرة الاستدلالية، والتي تنقسم بأنها عوامل ضيقة وغير قابلة لإعادة لاستخراج.

3- العامل الخاص أو النوعي: هو العامل الذي يختص بنوع واحد من أنواع السلوك الإنساني، ويوجد في متغير (اختبار) واحد فقط، أو عدة اختبارات تعكس جميعاً نفس المتغير المقاس، كاختبار الحساب، أو اختبار معاني الكلمات والتشابهات... إلخ. والتميز بين العوامل الثلاثة (العام والطائفي والنوعي) ليس تمييزاً قاطعاً، إنما يتوقف على عدد المتغيرات (الاختبارات) الخاضعة للتحليل، ومدى تباين أو تجانس استجابات أفراد العينة التي يجرى عليها التحليل العاملي، عدد أفراد العينة التي يجرى عليها التحليل العاملي، مدى تجانس أو تباين هذه الاختبارات في قياسها لما تقيس، حجم أو قيم معاملات الارتباطات البينية للاختبارات الخاضعة للتحليل العاملي.

بمعنى أن العامل الذي يظهر في مجموعة أقل عدداً من الاختبارات على أنه عامل عام، قد يظهر كعامل طائفي في مجموعة أكبر عدداً من الاختبارات، خاصة إذا مالت مجموعة منها إلى التجانس (الزيات، 1995).

2-11 مفهوم التحليل العاملي Factor Analysis Concept

التحليل العاملي (Factor Analysis) هو أسلوب إحصائي يستهدف تفسير معاملات الارتباطات الموجبة التي لها دلالة احصائية - بين مختلف المتغيرات، وبمعنى آخر فإن التحليل العاملي عملية رياضية تستهدف تبسيط الارتباطات بين مختلف المتغيرات الداخلة في التحليل وصولاً إلى العوامل المشتركة التي تصف العلاقة بين هذه المتغيرات وتفسيرها. ويعد التحليل العاملي منهجاً إحصائياً لتحليل بيانات متعددة ارتبطت فيما بينها بدرجات مختلفة من الارتباط الشخصي في صورة تصنيفات مستقلة قائمة على أسس نوعية للتصنيف، ويتولى الباحث فحص هذه الأسس التصنيفية واستشفاف ما بينها من خصائص مشتركة وفقاً للآطار النظري والمنطق العلمي الذي بدأ به، حيث يبدأ التحليل العاملي، بحساب الارتباطات بين عدد من المتغيرات مثل أ، ب، ج، د، هـ، و، مثل الذكاء، القلق، الانطواء، التحصيل، والاكتئاب مثلاً، ونحصل على مصفوفة من الارتباطات بين هذه المتغيرات لدى عينة ما، ثم نتقدم بعد ذلك لتحليل هذه المصفوفة الارتباطية تحليلاً عاملياً لنصل إلى أقل عدد ممكن من المحاور أو العوامل تمكنا من التعبير عن أكبر قدر من التباين بين هذه المتغيرات، ذلك أن توقفنا

عند فحص هذه المصفوفة الارتباطية التي تتكون من عشرة معاملات ارتباط لا يؤدي إلى فهم كامل للمجال المشترك فيما بينها جميعاً، حيث يبين كل معامل من معاملات الارتباط في المصفوفة علاقة بسيطة بين متغيرين فقط من متغيراتها دون أن ينبئ بأهمية أو دور هذه العلاقة بين هذين المتغيرين ومتغير ثالث، وعلى ذلك لا نستطيع عند هذا المستوى أن نصل لتقدير للعلاقة المشتركة بين ثلاثة متغيرات معاً أو بين متغيرات المصفوفة الخمس إذ أن حصولنا على معامل بالارتباط بين أ، ب قدره 0.7 ومعامل آخر بين ب، ج قدره 0.7 أيضاً لا يعني بالضرورة أن الارتباط بين أ، ج يساوي 0.7 كذلك فقد يكون ما هو مشترك بين أ، ب غير ما هو مشترك بين ب، ج، ولا تصلح العلاقة الثنائية بين ب وأي من المتغيرين أ، ج لتقدير العلاقة بينهما في معاملات الارتباط البسيطة (فرج، 1991).

يهدف أسلوب التحليل العاملي إلى تلخيص المتغيرات المتعددة في عدد أقل تسمى (عوامل) بحيث يكون لكل عامل من هذه العوامل دالة تربطه ببعض (أو كل) هذه المتغيرات. ويمكن من خلال هذه الدالة إعطاء تفسير لهذا العامل بحسب المتغيرات التي ترتبط معه بشكل قوي. ولقد نشأ هذا الأسلوب أساساً من أجل تحليل التجارب والمقاييس النفسية بحيث يمكن إرجاع مجموعة معينة من الاختبارات إلى عامل الذكاء وأخرى إلى عامل الذاكرة وهكذا، وإن كان هذا لا يعني أن هذا الأسلوب لا يستخدم في مجالات أخرى. وترتكز فكرة التحليل العاملي على استخلاص مجموعة من العوامل مرتبطة بالمتغيرات الأصلية، بحيث تفسر هذه العوامل أكبر نسبة ممكنة من التباين في المتغيرات الأصلية. ويمكن استخدام التحليل العاملي لتحويل مجموعة مرتبطة من المتغيرات إلى مجموعة أخرى مستقلة تربطها بالمجموعة الأولى علاقات خطية. وفي كل الأحوال تمثل العلاقة بين المتغيرات الأصلية والعوامل في شكل معادلات على النحو التالي:

$$F_1 = \alpha_{11}X_1 + \alpha_{12}X_2 + \dots + \alpha_{1n}X_n \quad \text{مثال على المعادلة}$$

أما المعادلات كاملة فهي كالتالي :

$$F_1 = \alpha_{11}X_1 + \alpha_{12}X_2 + \dots + \alpha_{1n}X_n$$

$$F_2 = \alpha_{21}X_1 + \alpha_{22}X_2 + \dots + \alpha_{2n}X_n$$

$$F_m = \alpha_{m1}X_1 + \alpha_{m2}X_2 + \dots + \alpha_{mn}X_n$$

3-11 أهمية التحليل العاملي وميادينه :

يمكن تطبيق أسلوب التحليل العاملي بنجاح في عدد كبير من الميادين العلمية، واختصار الوقت واجهد اللازمين للتحليل في العديد من الأبحاث، ويمكن إيجاز أهم هذه التطبيقات فيما يلي:

في مجال الإحصاء:

يُعتمد على التحليل العاملي في دراسة الارتباط والانحدار المتعدد بطريقة سريعة ودقيقة. فعلى سبيل المثال يستخدم التحليل العاملي لإثارة عدد من الفروض التي لها علاقة بالعوامل السببية، أو يستخدم لفحص المتغيرات قبل استخدامها في تحليل آخر، مثلاً: يمكن استخدام التحليل العاملي للبحث عن العلاقات الخطية المتعددة بين المتغيرات Multiconltnerity قبل تطبيق الانحدار المتعدد وتحويلها إلى عوامل مستقلة عن بعضها.

في مجال العلوم النفسية والاجتماعية والتربوية:

كُثر استخدام التحليل العاملي في هذا المجال، وذلك في تحليل النشاط العقلي المعرفي إلى قدراته المختلفة وتحليل النواحي المزاجية للشخصية إلى سماتها المتعددة وتحليل الاتجاهات والقيم الاجتماعية والميول المهنية.

في مجال بناء الاختبارات :

يعتمد بُناء الاختبارات الحديثة في دراسة مفردات الاختبارات على معرفة المكونات الرئيسية لظواهر التي تخضعها للقياس، ويُعد التحليل العاملي أدق وسيلة لمعرفة صدق هذه المكونات لقياس الظاهرة، وهو ما يسمى "بالصدق العاملي".

في مجال العلوم السياسية والتجارية :

حيث أن التحليل العاملي يقوم على الإيجاز الدقيق، فقد استخدم بنجاح كبير في دراسة الطواهر المعقدة التي تتأثر بعدد كبير من المؤثرات والعوامل المختلفة كالعلوم السياسية والإدارية، ودراسة العوامل المؤثرة في أسعار السلع والعملات وأجور العمال وما إلى ذلك. مجالات أخرى كثيرة مثل مجال العلوم الطبية، ومجال العلوم الطبيعية :

يستخلص من ذلك أن التحليل العاملي ليس وقفاً على علم النفس أو التربية فقط، ولكنه أسلوب علمي إحصائي من أساليب الدراسة التحليلية التي تهدف إلى التقسيم والتبويب والتصنيف لجميع القوى والمؤثرات الفعالة في ظاهرة معينة (باهي 2002م، غنيم 2000، المالكي 2000).

4-11 أهداف التحليل العاملي :

كما أن الهدف الأساسي من التحليل العاملي هو وصف علاقات التغير بين عدد كبير من المتغيرات بدلالة عدد قليل من المقادير غير المشاهدة التي تسمى العوامل، ويعتمد النموذج العاملي أساساً على الفكرة التالية: افتراض إمكانية جميع المتغيرات بناءً على معاملات الارتباط بينها، وهذا يعني أن جميع المتغيرات الموجودة في مجموعة معينة مرتبطة مع بعضها ارتباطاً قوياً، ولكن ارتباطها بمتغيرات المجموعات الأخرى ارتباط ضعيف، ومن الممكن أن نتصور هنا أن كل مجموعة من المتغيرات تمثل عاملاً واحداً، وهو المسئول عن الارتباط المشاهد بينها (باهي 2002م، غنيم 2000، المالكي 2000).

ويسعى أسلوب التحليل العاملي إلى استخلاص العوامل من المتغيرات بحيث:

1. يكون العامل الأول هو أكثرها ارتباطاً بالمتغيرات أو أكثرها تفسيراً للتباين المشترك بليه العامل الثاني وهكذا.
 2. أن يكون في كل عامل عدد غير قليل من المعاملات الصفرية.
 3. أن يسهل تفسير هذه العوامل على ضوء علاقاتها بالمتغيرات.
- من أهم أهداف العلم تنظيم الحقائق والمفاهيم تنظيمًا بوضع ما بينها من علاقات، أو تقسيمها على أساس ما بينها من أوجه التشابه والاختلاف والتحليل العاملي وسيلة من وسائل

التبسيط العلمي والتقسيم العلمي ويذكر "كاتل" (Cattell, 1952, 11) أن هدف المنهج العلمي اكتشاف الحقائق والعلاقة بين هذه الحقائق، ولأهداف عملية، واكتشاف القوانين التنبؤية، ويضيف أن التحليل العاملي منهج كلي يهدف إلى اكتشاف العموميات الأساسية، الوظيفية والعضوية، بدلا من أن ينوه البحث في عدد ضخم من المتغيرات التي تعد كالذرات، ولذلك يقترح "كاتل" أن يسمى بالتركيب العاملي أو على الأقل بتركيب المتغيرات. ومعنى أضيق يحدد "سولمون دياموند" (عبد الخالق، 1994: 99). أهداف التحليل العاملي بأنه تكوين الفروض واختبارها، وتحديد أصغر عدد من العوامل المحددة التي يمكن أن تفسر العلاقات التي نلاحظها بين عدد كبير من الظواهر الواقعية وإلى أي مدى يؤثر كل من هذه العوامل في كل متغير؟ أن أوضح وظيفة للتحليل العاملي تتمثل في خفض أو اختزال مكونات جداول الارتباطات إلى أقل عدد ممكن ليسهل تفسيرها.

لقد بين "أيزنك" (Eyzanck, 1953) أن للتحليل العاملي ثلاثة أهداف أساسية يروم تحقيقها، ويرتبط بهذه الأهداف ثلاث وجهات للنظر إلى طبيعة العوامل، وعدد كبير من طرق استخراج العوامل والتدوير، وهي الأهداف ذاتها لأي فرع من فروع الإحصاء وهي:

1. الوصف.
2. البرهنة على الفروض.
3. اقتراح فروض من البيانات الأولية.

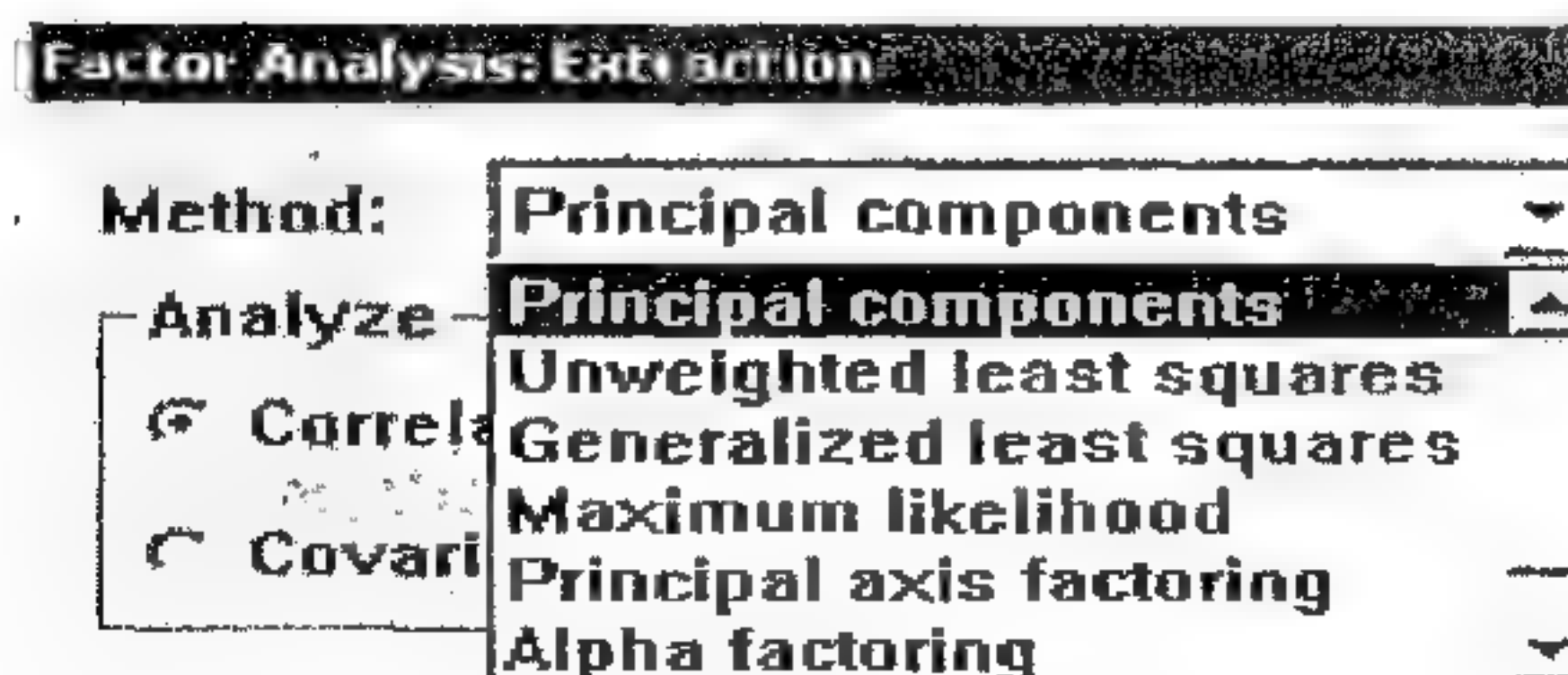
ومعظم علماء النفس يدركون هذه الاستخدامات الثلاثة للإحصاء، ولكن تظهر هذه المشكلة عندما تنطبق هذه الأهداف على التحليل العاملي، ويناقش "أيزنك" استخدامات التحليل العاملي على هذه المستويات الثلاثة، مع تعريف العامل في كل مستوى. فبالنسبة للهدف الأول فإن العامل إحصاء مختصر يهدف إلى اقتصاد في الوصف، ويصف علاقات مستقيمة بين مجموعة من المتغيرات، ولا يتضمن العامل تحديدا لأي معنى سيكولوجي أو أسباب، ولا يقترح فروضا أو يشتها، وقد وجد بعض علماء النفس وجهة النظر هذه جدا جدابة.

ويعتقد آخرون في عكس هذا الرأي، فيرون أن التحليل العاملي يقترح فروضا، وكلا نجح في هذه المهمة انتهت وظيفة الوصف ليصبح جزء من النظرية السيكلوجية من حيث هو

الاحصاء يختصر العلاقات بين مجموعة من المتغيرات، ويقترح علاقات سببية لم يسبق اكتشافها، وأن توليد الفروض ليس حكراً على التحليل العاملي، فهو يشبه في ذلك طرق الملاحظة والعمل الكلينيكي، إلا أن الآخرين يقللون عنه في درجة الدقة والصرامة. وقد يسهل تكوين الفروض في مجال تتوفر فيه ملاحظات كثيرة إلا أن إسهام التحليل العاملي يصبح مهماً جداً في المجالات الجديدة نسبياً، وذلك في الإسراع بتكوين فروض معقولة واستبعاد الفروض الضعيفة. ويتصل هذا الهدف بإثبات الفروض أو دحضها وبخاصة الفروض المتعلقة بتركيب الشخصية وتنظيمها كفروض الانمط والسمات، مما يصعب إثباته أو دحضه بالطرق غير العاملية. وبين مستوى اقتراح الفروض والتحقق منها رابطة متينة، وقد نجد النوعين من العوامل في دراسة واحدة. وحيث أن التحليل العاملي يهدف إلى تحقيق واحد أو أكثر من هذه الأهداف الهامة والجوهرية والتي تتسق مع أهداف العلم الأساسية، فقد أصبح التحليل العاملي منهجاً إحصائياً له أساس منطقي لا غنى عنه في عدد غير قليل من النظريات السيكولوجية وبالتحديد في مجال الشخصية التي تدعى عاملية (بدر الانصاري، 1997).

11-5 طرق التحليل العاملي :

تحدد الطرق الحسابية المستخدمة في التحليل العاملي كثيراً، فهناك الطريقة القطرية، والطريقة المركزة، والطريقة المركزة باستخدام متوسط الارتباطات، وطريقة المكونات الأساسية، ونوجزها فيما يلي:



1- الطريقة القطرية : Diagonal Method وتعد الطريقة القطرية من الطرق المباشرة

والسهلة في التحليل العاملي، ويمكن استخدامها إذا كان لدينا عدد قليل من المتغيرات وتؤدي إلى استخلاص أكبر عدد ممكن من العوامل وتتطلب هذه الطريقة معرفة سابقة ودقيقة بقيم شيوع

المتغيرات ، وبدون هذه المعرفة لا يمكن استخدامها، وتستمد الطريقة القطرية اسمها من كونها تقوم على استخدام القيم القطرية في المصفوفة الارتباطية مباشرة . وتبدأ الطريقة القطرية باستخلاص هذه القيمة بكاملها في العامل الأول، وبذلك يكون جذر هذه القيمة هو تشبع المتغير الأول على العامل الأول، ويطلق عليه اسم التشبع القطري وهكذا.

2- الطريقة المركزية : Centroid Method كانت الطريقة المركزية " لثريستون "

أكثر طرق التحليل العاملي استخداما وشيوعا إلى عهد قريب نظرا لسهولة حسابها فضلا عن استخلاص عدد قليل من العوامل العامة . غير أن هذه الطريقة تفتقر إلى عدد من المزايا الهامة ، أهمها أنها لا تستخلص الا قدرا محدودا من التباين الارتباطي ، تتحدد قيم الشيوع في المصفوفة الارتباطية وفق تقديرات غير دقيقة حيث تستخدم أقصى ارتباط بين المتغير وأي متغير في المصفوفة وهو إجراء يؤدي إلى خفض رتبة المصفوفة .

3- الطريقة المركزية باستخدام متوسط الارتباطات : Averoid Method لا تختلف

هذه الطريقة عن الطريقة المركزية المعتادة إلا في استخدامها تقدير الشيوع عبارة عن متوسط ارتباطات المتغير ببقية المتغيرات في المصفوفة ثم حساب العوامل بعد وضع المتوسط الخاص بارتباطات كل متغير في خليته القطرية ولهذا السبب يطلق على هذا الأسلوب اسم الطريقة المركزية باستخدام المتوسطات. غير أن هذه الطريقة لا توفر نفس الدقة التي تجدها في الطريقة المركزية التامة ، إذ تؤدي إلى خفض محدود في نسبة التباين التي تعبر عنها العوامل الناتجة. غير أن هذه الطريقة تبدو مفيدة في حالة وجود عدد كبير من المتغيرات دون توفر وسائل آلية لإجراء العمليات الحسابية.

4- أسلوب تحليل المركبات الأساسية Principal Components Analysis

(P.C.A.)

طريقة المكونات الأساسية : Principal Componants تعد طريقة المكونات

الأساسية التي وضعها "هوتلينج Hottelling" عام 1933 من أكثر طرق التحليل العاملي دقة وشيوعاً في بحوث الشخصية ، ولهذه الطريقة مزايا عدة منها أنها تؤدي إلى تشبعات دقيقة. وكذلك " فإن كل عامل يستخرج أقصى كمية من التباين (أي أن مجموع مربعات تشبعات العامل تصل إلى أقصى درجة بالنسبة لكل عامل) ، وتؤدي إلى أقل قدر ممكن من البواقي، كما

أن المصفوفة الارتباطية تختزل إلى أقل عدد من العوامل المتعامدة (غير المرتبطة) ولم تلق طريقة المكونات الأساسية في البداية قبولا كبيرا بين الباحثين نظرا لحاجتها إلى وقت حسابات طويل لإتمامها ولذا كان من المستحيل استخدامها يدويا في حالة المصفوفات الكبيرة ، ولكن بعد الاعتماد على الآلات الحاسبة الالكترونية ذات السرعة الفائقة والدقة الشديدة وطاقة التخزين الكبيرة ، أصبحت هذه الطريقة الآن من بين أكثر الطرق شيوعا نظرا لدقة نتائجها بالمقارنة ببقية الطرق.

يعتمد أسلوب تحليل المركبات الأساسية (P.C.A.) بصفة أساسية على تفسير وتحليل مجموعة التغيرات والتباينات بين البيانات من خلال مجموعة صغيرة من التوليفات الخطية في المتغيرات الأساسية. ومن ثم فإن الهدف الأساسي لهذا الأسلوب التحليلي هو تفسير البيانات ومعرفة مدى اختلافها وأسباب هذا الاختلاف، وكذلك التعامل مع البيانات بصورة مختصرة Data Reduction من خلال أقل عدد ممكن من العلاقات الخطية والتي تفسر في مجملها أكبر جزء ممكن من الاختلافات والتباينات بينها.

بفرض أن لدينا n من المتغيرات $X = (X_{1j}, \dots, X_{pj}, j = 1, \dots, N)$ مجتمع حجمه N من المشاهدات لـ n من المتغيرات في الصورة التالية:

$$X = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1N} \\ X_{21} & X_{22} & \dots & X_{2N} \\ \vdots & \vdots & \dots & \vdots \\ X_{n1} & X_{n2} & \dots & X_{nN} \end{bmatrix}$$

حيث $n > p$ وإذا كانت مصفوفة التباينات والتغايرات هي S ، فإن المشكلة تكمن في إيجاد العلاقات أو التوليفات الخطية⁽¹⁾

⁽¹⁾ Sturt, M., (1982), "A Geometric Approach to Principal Components Analysis" The American Statistician, 36, 365 - 367.

$$\begin{aligned}
 y_1 &= a_{11}X_1 + a_{21}X_2 + \dots + a_{p1}X_p = \underline{a}'_1 \underline{X} \\
 y_2 &= a_{12}X_1 + a_{22}X_2 + \dots + a_{p2}X_p = \underline{a}'_2 \underline{X} \\
 &\text{M} \quad \text{M} \quad \text{M} \quad \text{M} \quad \text{M} \quad \text{M} \\
 y_p &= a_{1p}X_1 + a_{2p}X_2 + \dots + a_{pp}X_p = \underline{a}'_p \underline{X}
 \end{aligned}
 \tag{3-1}$$

حيث

$$\text{var}(y_1) = \underline{a}'_1 \Sigma \underline{a}_1 \quad 1 = 1, 2, \dots, p \tag{3-2}$$

$$\text{cov}(y_1, y_k) = \underline{a}'_1 \Sigma \underline{a}_k \quad 1, k = 1, 2, \dots, p \tag{3-3}$$

وتكون المركبات الأساسية هي تلك التوليفات الخطية غير المرتبطة (Uncorrelated linear combination) y_1, y_2, \dots, y_p التي تجعل التباينات (3-2) أكبر ما يمكن. ومن ثم يكون المركب الأساسي الأول هو التوليفة الخطية ذات أعلى تباين أي التوليفة الخطية التي تحقق القيمة العظمى للتباين $\text{var}(y_1) = \underline{a}'_1 \Sigma \underline{a}_1$ بشرط أن $\underline{a}'_1 \underline{a}_1 = 1$. ويكون المركب الأساسي الثاني هو التوليفة الخطية ذات أعلى التباينات المتبقية في البيانات، أي التوليفة الخطية التي تحقق القيمة العظمى للتباين.

$$\text{var}(y_2) = \underline{a}'_2 \Sigma \underline{a}_2 \tag{3-4}$$

تحت شروط

$$\underline{a}'_2 \underline{a}_1 = 0$$

$$\text{cov}(\underline{a}'_1 \underline{X}, \underline{a}'_2 \underline{X}) = 0$$

وهكذا يكون المركب الأساسي رقم 1 هو التوليفة الخطية $\underline{a}'_1 \underline{X}$ التي تحقق القيمة العظمى لتباينات المتبقية أي العلاقة الخطية $Y_1 = \underline{a}'_1 \underline{X}$ التي تحقق القيمة العظمى لتباين.

$$\text{var}(y_1) = \underline{a}'_1 \Sigma \underline{a}_1 \tag{3-5}$$

تحت شروط

$$\underline{a}'_1 \underline{a}_1 = 1$$

$$\text{cov}(\underline{a}'_1 \underline{X}, \underline{a}'_k \underline{X}) = 0 \quad \text{for} \quad k < 1$$

وهكذا حتى نحصل على عدد من التوليفات يستحوذ على أكبر قدر ممكن من التباينات.

ويمكن التوصل إلى هذه العلاقات باستخدام فكرة مضروب لاجرانج حيث أنه بالنسبة للمركب الأساسي الأول فإن :

$$\begin{aligned}\phi_1 &= \text{var}(y_1) - \lambda_1(a_1' a_1 - 1) \\ &= a_1' \Sigma a_1 - \lambda_1(a_1' a_1 - 1)\end{aligned}\quad (3-6)$$

حيث l_1 هي مضروب لاجرانج للمركب الأساسي الأول . وبإجراء التفاضلات الجزئية فإن :

$$\begin{aligned}\frac{\partial \phi_1}{\partial a_1} &= 2 \Sigma a_1 - 2 \lambda_1 a_1 = 0 \\ \frac{\partial \phi_1}{\partial \lambda_1} &= a_1' a_1 - 1 = 0\end{aligned}\quad (3-7)$$

ومن ثم نحصل على الآتي :

$$(S - l_1 I) a_1 = 0 \quad (3-8)$$

وللوصول إلى حل العلاقة (3-8) بشرط أن $a_1' a_1 = 1$ من

$$|S - l_1 I| = 0 \quad (3-9)$$

ومن ثم يمكن التوصل إلى كل من الجذر الكامن الأول (l_1 eigenvalue) والمتجه الكامن الأول (a_1 eigenvector) والنماظر للجذر l_1 .

أما بالنسبة للمركب الأساسي الثاني فإن $y_2 = a_2' X$

$$\begin{aligned}\phi_2 &= \text{var}(y_2) - \lambda_2(a_2' a_2 - 1) - \mu(a_2' a_2) \\ &= a_2' \Sigma a_2 - \lambda_2(a_2' a_2 - 1) - \mu(a_2' a_2)\end{aligned}\quad (3-10)$$

حيث l_2 هو مضروب لاجرانج للمركب الأساسي الثاني وبإجراء التفاضلات الجزئية

فإن :

$$\begin{aligned}\frac{\partial \phi_2}{\partial a_2} &= 2 \sum a_2 - 2 \lambda_2 a_2 - \mu a_1 = 0 \\ \frac{\partial \phi_2}{\partial \lambda_2} &= a_2' a_2 - 1 = 0 \\ \frac{\partial \phi_2}{\partial \mu} &= a_1' a_2 = 0\end{aligned}\tag{3-11}$$

ومن ثم نحصل على العلاقة التالية :

$$(S - l_2 I) a_2 = 0\tag{3-12}$$

ولوصول إلى حل العلاقة (3-12) بشرط $a_1' a_2 = 0$ ، $a_2' a_2 = 1$ من

$$|S - l_2 I| = 0\tag{3-13}$$

ومن ثم يمكن التوصل إلى كل من الجذر الكامن الثاني (l_2 eigenvalue) والمتجه الكامن الثاني (a_2 eigenvector) والمناظر للجذر l_2 .

وهكذا يمكن الحصول على المركب الأساسي رقم 1 ، كما أن a_1 هو المتجه الكامن المناظر له، ومن ثم يكون لدينا عدد من الجذور الكامنة بحيث أن :

$$l_1 \geq l_2 \geq l_3 \geq \dots \geq l_p$$

وينظر كل جذر منها متجه كامنا يمثل ثوابت العلاقة الخطية الجديدة، مع ملاحظة أنه على الرغم من أنه يكون لدينا عدد من التوليفات الخطية مساويا p ، فإن معظم التباينات والاختلافات بين البيانات والملاحظات ترجع إلى عدد أقل من التوليفات وليكن m توليفة (m مركب أساسي) حيث أن معظم المعلومات يرجع تفسيرها إلى هذا العدد الصغير من مركبات، ومن ثم يمكن الاستعاضة عن p بمتغير بعدد m مركبة أساسية لتفسير الظاهرة محل الدراسة، وبالتالي يمكن التعامل مع مجموعة من المعلومات والبيانات حجمها $n \times m$ بدلا من $n \times p$.
والجدير بالذكر أن من صفات أسلوب تحليل المركبات الأساسية صفة التغير والتنوع للعلاقات الخطية الناتجة إلا أنه يختص بخاصيتين وهما⁽¹⁾ :

¹ Jackson J E., (1980): Principal components and factor analysis. Part I - Principal Components", Journal of Quality Technology, 12, 201-213.

$$(1) \quad |S| = |I|$$

$$(2) \quad \text{tr } S = \text{tr } I$$
(3-15)

حيث :

$$\text{tr } S = S_{11} + S_{22} + \dots + S_{pp} \quad \text{بمجموع التباينات}$$

$$\text{tr } I = I_1 + I_2 + \dots + I_p \quad \text{بمجموع الجذور الكامنة}$$

أي أن مجموع التباينات تساوي مجموع الجذور الكامنة eigenvalues ، كما أن محدد مصفوفة التباينات تساوي محدد الجذور الكامنة.

وبالتالي يمكن قياس درجة الأهمية النسبية لكل من مركب اساسي لتفسير سلوك المتغيرات الأصلية في تشكيل الظاهرة باقل خطأ ممكن للمعلومات في النظام كالتالي :

$$RI_i = \frac{\lambda_i}{\text{tr } \Sigma} \quad (3-16)$$

$$\frac{\text{الجذر رقم I الكامن}}{\text{بمجموع التباينات}} = \text{أي أن الأهمية النسبية } (RI_i) \text{ للمركبة الأساسية I مثلا}$$

حيث تمثل الأهمية النسبية هنا إجمالى التباينات بين المشاهدات المفسرة باستخدام المركبة الأساسية رقم I يلاحظ أنه في أغلب الأحيان أن حوالى من 80% إلى 90% من إجمالى التباينات لعدد كبير من المتغيرات p يمكن رجوعه إلى المركبة الأساسية الأولى أو المركبتين الأساسيتين الأوليين أو المركبات الثلاث الأولى ، ومن ثم يمكن الاستعاضة عن العدد الأصلى p بهذه المركبات الثلاث دون فقد يذكر للمعلومات المتاحة.

طريقة المكونات الاساسية Principal components

يشير صفوت فرج (1980) ان طريق المكونات الاساسية ل هوتلنج من اكثر طرق التحليل العاملي دقة ومميزات، غير ان الكثيرين من الباحثين كانوا يحجمون عن استخدامها لما تتطلبه من اجراءات طويلة وعمليات حساية متعددة الا انه ازاء التقدم العلمى الراهن في استخدام الحاسبات الالكترونية الحديثة والفائقة السرعة مثال في البحوث النفسية اصبح من غير

المستطاع مقاومة اغراء استخدام هذه الطرق الدقيقة SPSS , SAS ، وطريقة المكونات الأساسية لا تفترض تسلسل التباين النوعي في شكل عوامل نوعية ويدمج هذا التباين في هذه الطريقة في التباين العام مكونا فئات تصنيفية كبرى تتضمن نسبة ضئيلة من هذا التباين النوعي لا تظهر واضحة في العوامل المبكرة الاستخلاص عامليا والتي تعد ذات اهمية كبيرة في هذا الاسلوب.

يضاف الى ذلك ميزة رئيسية في المكونات الأساسية هي ان كل عامل فيها يستخلص اقصى تباين ممكن ، بمعنى ان مجموع المربعات يصل الى اقصى حدوده في كل عامل وعلى ذلك تتلخص المصفوفة الارتباطية في اقلا عدد من العوامل المتعامدة.

وهذا معناه ان اسلوب المكونات الأساسية يتميز بقدرة على الوصول الى حل يتفق مع محك أو في مربعات للمصفوفة الارتباطية وهو احد المحكات الرياضية التي تلاقى قبولا وضحا في مجال الاساليب التلخيصية للعلاقات بين المتغيرات تهتم بعملية شرح وتفسير بناء التباين (فرج، 1980) . وطريقة المكونات الأساسية Principal components وبناء التباين التلازمي أو

التغاير المزدوج Variance structure

من خلال الارتباطات الخطية القليلة Covariance structure بالمتغيرات الاصلية وموضوعات العامة هي: Linear combination اختزال او تحويل البيانات Data reduction الشرح او التفسير Interpretation (Johenson, & Wichern , 1992 : 256)

1. التدوير المتعامد والمائل

هناك نوعان من التدوير تبعاً للزاوية التي تفصل بين المحاور المرجعية وهما المتعامد والمائل Orthogonal ، ففي التدوير المتعامد تدار العوامل معاً (اثنين منهما مثلاً) مع الاحتفاظ بالتعامد oblique (90 درجة) اما التدوير المائل ففئة تدار المحاور دون احتفاظ بالتعامد ، فتترك لتتخذ الميل الملائم لها .

وفي هذا الصدد يذكر محمود منسى (1993) ان جميع العوامل المستخلصة من التحليل العاملى المباشر ، تحتاج الى ابراز هويتها بطريقة اوضح لانه يصعب تفسيرها سيكولوجيا وحيث ان هدف تدوير المحاور (المتعامد والمائل على حد سواء) هو التوصل الى البناء البسيط الا ان

التدوير المتعامد يهدف الى تحقيق هذا الهدف في ضوء فكرة الاستقلال بين العوامل وعدم الارتباط (جتا $90 = 0$) (منسى ، 1993 : 70)

والعوامل المتعامدة غير مرتبطة معا ، اى ان معاملات الارتباط بينها تساوى الصفر ، اد تصنف العوامل الاختبارات أو المتغيرات الى فئات غير مرتبطة ، وهكذا يصبح التقسيم حاد غير مستداخل ، اما العوامل المائلة فهي بينها ارتباط اى انها عوامل متداخلة وبفضل بعض المحلين استخراج عوامل متعامدة غير مرتبطة في حين يهتم آخرون باستخلاص المائل ، ويهدف تدوير المحاور الى تحقيق ما يسمية ثرستون البناء البسيط (عبد الخالق ، 1987 : 116).

بعض طرق التدوير المتعامد

- الكوارتيماكس
- البارماكس
- الماكسبلان

بعض طرق التدوير المائل

- الكوارتيمن
- الاوبلمن
- الكوفاريمين
- البروماكس

محكات التوقف عن استخلاص العوامل : (تدوير العوامل)

1- محك تاكر **Tuker Phi** : وهو يقوم على استخدام معامل فاي ويعتمد على مبدأ انه اذا لم يوجد نقص ذي دلالة في حجم قيم البواقي من مصفوفة الى اخرى تليها فان العامل السذي استخلص يكون ذا دلالة وما يتبقى ليس الا بواقي لا اهمية لها. ويلاحظ ان هذه الطريقة تصلح بالنسبة للطريقة المركزية لثرستون ونعتمد صحة هذا المحك جزئيا على صحة اجراءات عكس الاشارات في المصفوفة كما ان عملية العكس تؤدي الى ارتفاع في القيم الايجابية مما يترتب عليه ان تتجاوز قيمة فاي قيمة $(1-n) / (1+n)$.

2- قاعدة همفري : تعتمد هذه القاعدة على حجم العينة الاصلية التي حسبت الارتباطات بين متغيراتها وثانيا على فكرة ان تشبع متغيرين فقط دون المصفوفة كلها كافيين تماما لتقرير وجود عامل عام .

3- محك كومب : يطبق هذا المحك فقط على المصفوفات التي تحتوي على قيم موجبة او صفرية ويسمح بالقيم السالبة الصغيرة التي لا تختلف اختلافا واضحا عن الصفر وبذلك يعتمد هذا الاسلوب على نمط البواقي في المصفوفة اكثر من اعتماده على حجمها او دلالاتها حيث انه يفترض انه في حالة وجود عوامل ذات دلالة مرتفعة لم تستخلص بعد وليس مجرد تباین خطأ المصفوفة فعلياً ان لا نتوقع قيم سالبة اكثر في مصفوفة البواقي بعد العكس مما يتوقع بحكم الصدفة في مصفوفة ناتجة عن ارتباطات ايجابية .

4- محك كايزر : يعتمد هذا المحك على حجم التباين الذي يعبر عنه العامل ، وعلى ذلك فان هذا المحك يتطلب مراجعة الجذر الكامن للعوامل الناتجة ، وعلى ان تقبل العوامل التي يزيد جذرها الكامن عن الواحد الصحيح ، وتعد عوامل عامة . وهو محك قد يكون صالحا ومناسبا لطريقة المكونات الاساسية لهوتلنج على وجه الخصوص . ويذكر عبد الخالق (1994) ان العوامل الدالة في هذه الطريقة هي العوامل التي يساوي او يزيد جذرها الكامن على واحد صحيح أي ان التباين الذي يستوعبه كل عامل (مجموع مربعات التشبعات على كل عامل) ≥ 1 .. بشرط ان يكون قد وضع في الخلايا القطرية واحد صحيح ويذكر وايت وزملاؤه ان هذا المعيار تتطابق نتائجه مع معايير اخرى .

5- محك كاتل : تؤدي خطوات استخلاص العوامل من المصفوفة الارتباطية الى انتاج العوامل الاكثر عمومية اولا في كل الاساليب العاملة بلا استثناء ، ثم تبدأ العوامل الخاصة او التباين الوعي في الظهور . وفي طريقة المكونات الاساسية لاتفرق بين عوامل عامة واخرى غير عامة يفترض ايضا ان حجم التباين الوعي الذي يتسرب الى العوامل الناتجة يتزايد في العوامل الاخيرة ويبدأ في فرض صورة تقلل من اهمية المصفوفة العاملة ويتطلب الامر في هذه الحالة تحديد العدد الامثل من العوامل قبل ان تؤدي ظهور التباينات الخاصة الى احداث خلل في مصفوفة العوامل ويقترح كاتل هنا محكا بسيطا يطلق عليه اسم البقايا المبعثرة وذلك بان يقوم برسم محورين متعامدين ، افقي نضع عليه عدد العوامل في تحليلنا (الذي انتج فيه عددا كبيرا من

العوامل) ويقسم المحاور الراسي وفقا لوحداث منتظمة معبرة عن الجذر الكامن المستخلص للعوامل المختلفة.

وسنلاحظ بعد اتمام رصد عواملنا وجذورها الكامنة ، ان حجم الجذر يتناقض بشكل كبير في العوامل الاولى الى ان يصل الى نقطة معينة هي غالبا حول جذر كامن واحد صحيح ثم يبدأ حجم الجذر في التناقص بصورة ضئيلة بحيث يستوي فيها الخط البياني مع الخط الافقي . وإذا افترضنا أن النقطة التي سنتوقف لديها في قبولنا للعوامل هي عند العامل الرابع على سبيل المثال فان الفرق لن يكون كبيرا في الواقع بين ما يقدمه محك كاتل وبين ما يقدمه محك كايزر الذي يتطلب التوقف عند العامل الثالث هذا على سبيل المثال .

وتبقى لطريقة كايزر ميزتها في هذه الحالة في كونها لا تتطلب استخلاص عدد كبير من العوامل ثم رصدها في الشكل البياني للتعرف على نقطة توقف التناقص واستواء الخط ، حيث يمكن حساب الجذر الكامن لكل عامل بطريقة كايزر قبل استخلاص العامل التالي مما يوفر جهدا لا مبرر له (صفوت فرج ، 1991 : 246) .

6- محك مويزز : يقوم هذا المحك على تفرطح التباين الكلي للعوامل المتتالية .

7- محك بيرت وبانكز : ويمكن عن طريق هذا المحك تحديد العوامل ذات الدلالة المنخفضة عن طريق تحديد الخطأ المعياري للنشبع الصفري ، ومقارنة عدد تشبعات العامل أو مضاعفات هذا العدد التي يزيد مقدارها عن الخطأ المعياري .

11-6 بعض مفاهيم التحليل العاملي:

1- درجة الشيوخ Communalty:

تعرف درجة شيوخ المتغير بإسهامات هذا المتغير في جميع العوامل ويقاس بمجموع مربعات معاملات هذا المتغير في العوامل المختلفة، فمثلا تقاس درجة شيوخ المتغير رقم (J) على النحو التالي:

$$C_j = \sum_{i=1}^m \alpha_{ji}^2$$

درجة التشبع Loading :

$$\alpha_{ij}$$

يعرف المعامل

بمعامل تحميل (أو تشبع) المتغير i على العامل j كما يعبر عن مدى ارتباط العامل بالمتغير. ويلاحظ أن مجموع مربعات درجات التشبع لكل عامل تسمى الجذر الكامن وتعبر عن أهمية هذا العامل في تفسير الاختلافات في المتغيرات، كما يعبر مجموع الجذور الكامنة عن التباين الذي أمكن تفسيره من خلال العوامل، ونسبته إلى عدد المتغيرات نحصل على نسبة التباين العملية هذه؛ لأنه يتم تحويل المشاهدات إلى قيم معيارية ويكون تباين كل متغير الواحد الصحيح.

2- معامل الارتباط كعلاقة بين الاختبار (1) والاختبار (2) يساوي=

حاصل ضرب درجة تشبع الاختبار (1) بعامل معين (أ) \times درجة تشبع الاختبار (2) بعامل معين (أ).

$$\text{أي أن } 1.2 = \text{ش}^1 \times \text{ش}^2$$

وقياسا على ذلك فإن معامل الارتباط بين الاختبار (1) ونفسه = (ش¹)²

3- التباين Variance:

يحسن الاعتماد في التحليل العاملي على الدرجات المعيارية Z - Score وهي تعني توحيد أساس الدرجات على المتغيرات المختلفة بحيث تصبح وحدة الدرجة الخاصة بالفرد على المتغير واحد صحيح أو درجة أي فرد عبارة عن نسبة من هذا الواحد الصحيح وبهذا نلاحظ أن تباين المتغير الواحد في أي تحليل عاملي هو:

$$\text{التباين الكلي} = \text{تباين العامل العام} + \text{تباين العامل النوعي} + \text{تباين الخطأ}$$

حيث أن: تباين المتغير العام هو: مربع تشبع المتغير أو مربع ارتباطه بالعامل وهو هنا تباين عام يشترك به المتغير مع تباينات لمتغيرات أخرى بما يؤدي إلى استخلاص عامل عام تباين المتغير النوعي هو: مربع تشبع المتغير أو مربع ارتباطه بالعامل وهو قدر من التباين الذي يعبر به المتغير الواحد عن نوعية أدائه ويظهر على عامل دون أن يظهر معه تباين لمتغيرات أخرى.

تباين الخطأ هو :الجزء الذي لا يستخلص في شكل عوامل ويبقى في المصفوفة الارتباطية بعد استخلاص العوامل على شكل بقايا ويعود إلى عدد من الأسباب وهي:

أ- أخطاء القياس : ويقصد بها استخدام الأدوات منخفضة الثبات أو استخدام مقاييس غير متجانسة البنود ، أو تأثير بعض المتغيرات الأخرى فكل هذا يؤثر على نتائج التحليل العاملي .

ب- أخطاء التجربة : والتي تتمثل في عدم الضبط الدقيق للمتغيرات بالبحث .

ج- أخطاء الدقة : والتي تتمثل في عدم إحكام جلسة الاختبار أو طريقة تقديم التعليم أو أسلوب تصحيح الاختبارات .

4- العلاقة بين الثبات والشيوع :

أن معامل الثبات يعبر عن الحجم الحقيقي لتباين المتغير أي بعد استبعاد تباين الخطأ وأنا ننظر إلى قيم الشيوع للمتغير في مصفوفة عاملية باعتبارها معامل ثبات لهذا المتغير حيث تمثل قيم الشيوع في هذه الحالة هذا التباين الحقيقي الذي استخلص معبرا عن تباينات مختلفة يشترك فيها المتغير مع غيره من المتغيرات طالما بقي تباين الخطأ في مصفوفة البواقي معبرا بدوره عن الجزء من التباين الكلي الذي لا يشترك فيه الاختبار مع غيره من المتغيرات نتيجة لأخطاء القياس أو أخطاء التجريب.

5- الجذر الكامن: Eigen Value

يعرف مجموع مربعات تشبعات كل المتغيرات على كل عامل على حدة من عوامل المصفوفة باسم الجذر الكامن للعامل وهو تعبير يستخدم في جبر المصفوفات ويلاحظ بالنسبة لأي مصفوفة عاملية أن الجذر الكامن يتناقص تدريجيا عن العامل الآخر، فالعوامل الأولى ذات جذر كامن أكبر من العوامل المتأخرة الاستخلاص ، ذلك أن خطوات حساب العوامل تؤدي إلى استخلاص أقصى تباين مشترك بين المتغيرات في كل مرة على التوالي وبطرح مصفوفة الناتج من المصفوفة الارتباطية يتبقى حجم اصغر من التباين المشترك بين المتغيرات يستخلص في عامل جديد ذي جذر كامن اصغر من سابقه. وسيكون مجموع قيم الشيوع للمتغيرات يساوي تماما

بمجموع الجذور الكامنة لعوامل المصفوفة ، بمعنى آخر أن مجموع مربعات الصفوف ((أي قيم الشيوخ)) = مجموع مربعات الأعمدة ((أي الجذور الكامنة)).

ولما كان التباين الذي يساهم به متغير بقيمة المعيارية يساوي (1) فإن أي مكون جذره الكامن اقل من واحد لا يكون له أي أهمية تذكر ، ومعنى ذلك أن المكونات أو العوامل التي تكون قيمة الجذر الكامن لكل منها واحد أو أكثر هي التي تعتمد وتعتبر ذات دلالة معنوية

6- حجم التباين العاملي ونسبة التباين العاملي :

حجم التباين العاملي هو مجموع قيم الشيوخ أو مجموع الجذور الكامنة ، أما نسبة التباين العاملي للمصفوفة عبارة عن :

$$\text{مجموع الجذور الكامنة للعوامل} \times 100$$

التباين الارتباطي

والتباين الارتباطي يساوي عدد المتغيرات التي تدخل في التحليل العاملي والجذر الكامن يعكس مقدار التباين العام عن طريق العدد النسبي من العوامل.

11-7 استخدام برمجية SPSS في حل المسائل.

إذا كان لدينا استبانة تتكون من 23 سؤالاً حول البرنامج الاحصائي SPSS وهي

كالتالي:

مؤلف	مؤلف	مؤلف	مؤلف	مؤلف	استبانة حول الاحصاء والبرنامج SPSS
مؤلف	مؤلف	مؤلف	مؤلف	مؤلف	1. الاحصاء يجعلني أبكي.
مؤلف	مؤلف	مؤلف	مؤلف	مؤلف	2. يعتقد أصدقائي أنني غبي لأنني لم ابدأ بالتعامل مع البرنامج إحصائي للعلوم الاجتماعية.
مؤلف	مؤلف	مؤلف	مؤلف	مؤلف	3. الانحرافات المعيارية تثيرني.
مؤلف	مؤلف	مؤلف	مؤلف	مؤلف	4. أحلم بأن يرسون يهاجمني في معاملات الارتباط.
مؤلف	مؤلف	مؤلف	مؤلف	مؤلف	5. أنا لا أنهم الإحصاء.
مؤلف	مؤلف	مؤلف	مؤلف	مؤلف	6. ليس لدي الكثير من الخبرة في الحاسوب.
مؤلف	مؤلف	مؤلف	مؤلف	مؤلف	7. جميع الحواسيب تكن الكراهية لي.
مؤلف	مؤلف	مؤلف	مؤلف	مؤلف	8. أنا لست جيداً في الرياضيات.

					9. أصدقائي في الإحصاء أفضل مني.
					10. الكمبيوتر مقيدة فقط في اللعب.
					11. فعلت شيئاً في الرياضيات في المدرسة.
					12. يحاول الناس القول بأن SPSS يجعل الإحصاء أسير على الفهم ولكنه بالنسبة لي ليس كذلك.
					13. بسبب استخدام الكمبيوتر إنني أخشى أن يسبب ضرراً لا يمكن إصلاحه.
					14. الكمبيوتر يحتاج إلى عقول وأنا أشعر بأني على غير ما يرام عند استخدامها.
					15. أشعر بأن الكمبيوتر يرفضني.
					16. أنا أبكي صراحة عند ذكر مقاييس الزعة المركزية.
					17. أنا تصيبي حالة غيوبة كلما أرى المعادلات.
					18. تنحطم برامج الإحصائي للعلوم الاجتماعية دائماً عندما أحاول استخدامها.
					19. الجميع ينظر لي عندما أقوم باستخدام البرنامج الإحصائي للعلوم الاجتماعية.
					20. لا يمكنني النوم عندما أفكر في المتجهات الكامنة.
					21. أنا أشعر بالسوء عندما أفكر بالتوزيع الطبيعي.
					22. أصدقائي في البرنامج الإحصائي للعلوم الاجتماعية على نحو أفضل مما أنا عليه.
					23. إذا كنت جيداً في الإحصاء فإن أصدقائي سوف يفكرون إنني أنا الطالب الذي يذاكر كثيراً.

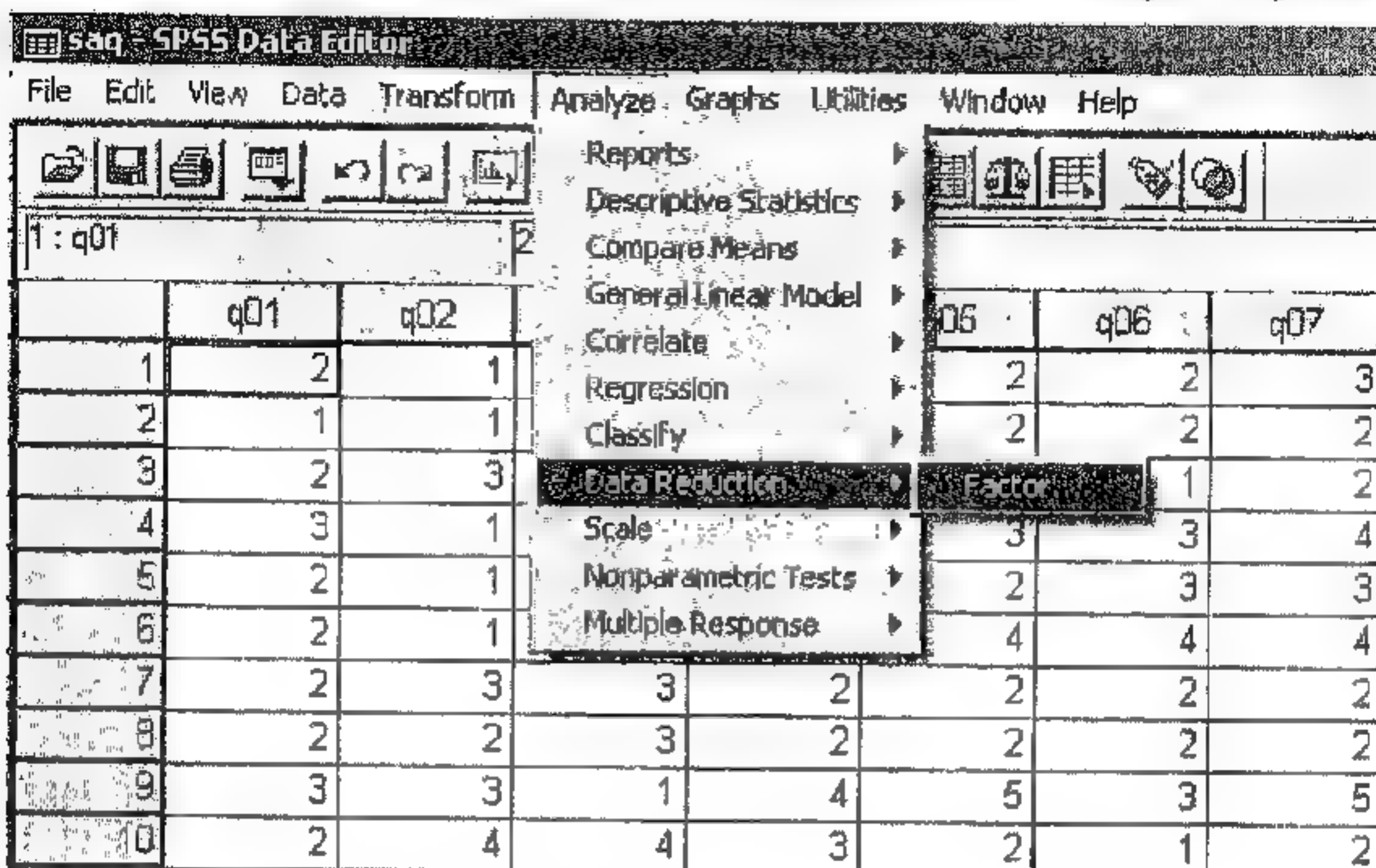
1. توصيف المتغيرات:

	Name	Type	Width	Decimal	Label	Values	Missing	Columns	Align	Measure
1	q01	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
2	q02	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
3	q03	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
4	q04	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
5	q05	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
6	q06	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
7	q07	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
8	q08	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
9	q09	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
10	q10	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
11	q11	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
12	q12	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
13	q13	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
14	q14	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
15	q15	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
16	q16	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
17	q17	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
18	q18	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
19	q19	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
20	q20	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
21	q21	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal
22	q22	Numeric	1	0		{1, Strongl	9	8	Right	Ordinal

2. إدخال البيانات:

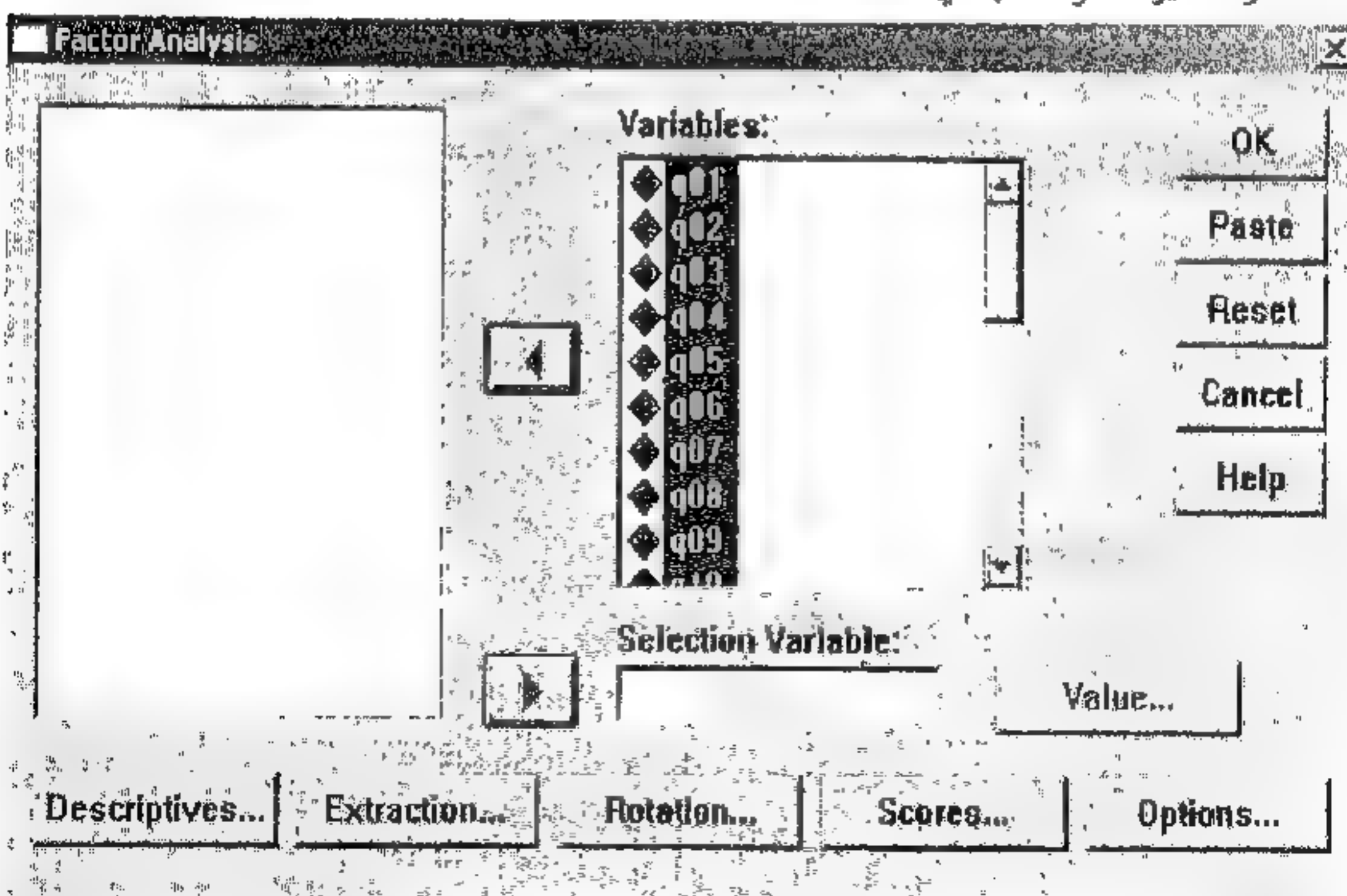
	q01	q02	q03	q04	q05	q06	q07	q08	q09	q10	q11	q12	q13	q14	q15	q16	q17	q18	q19	q20	q21	q22	q23
1	2	1	4	2	2	2	3	1	1	2	1	2	2	2	2	1	1	2	3	2	2	2	5
2	1	1	4	3	2	2	2	3	5	2	2	3	1	3	4	3	2	2	3	4	4	4	2
3	2	3	2	2	4	1	2	2	2	2	3	3	2	4	2	3	2	3	1	4	3	2	2
4	1	1	1	4	3	3	4	2	2	4	2	2	2	2	2	3	2	4	2	4	4	4	2
5	2	1	3	3	2	3	3	2	4	3	2	3	3	2	2	2	2	3	3	4	4	4	4
6	2	1	3	2	4	4	4	2	4	3	2	4	3	3	5	2	3	5	1	5	3	1	4
7	2	3	3	2	2	2	2	2	3	2	2	2	2	2	2	2	2	2	3	2	2	4	4
8	2	2	3	2	2	2	2	2	4	2	2	3	2	2	3	2	2	2	4	3	2	4	4
9	1	3	1	4	5	3	5	5	3	3	5	5	5	5	5	5	5	5	2	5	5	2	3
10	2	4	4	3	2	1	2	2	3	2	2	3	2	1	2	3	2	3	3	3	4	4	4
11	2	1	5	4	2	1	2	2	5	2	1	3	1	2	1	2	2	2	5	3	2	5	5
12	2	1	3	3	4	3	3	1	3	2	2	3	2	2	3	3	2	2	3	4	3	4	4
13	3	1	3	4	3	2	3	3	2	3	3	4	4	4	4	4	3	3	2	4	4	3	4
14	2	2	1	2	2	2	2	2	2	3	2	4	2	4	4	4	2	4	1	5	5	3	1
15	2	2	3	4	2	2	3	2	2	3	2	3	2	3	3	4	2	3	3	4	4	4	4
16	3	1	2	2	2	2	2	2	2	3	2	3	2	3	2	3	2	3	2	3	2	3	4
17	1	2	5	2	1	1	1	1	4	1	1	2	1	1	1	2	2	1	4	2	1	4	4
18	2	2	3	3	3	4	3	2	5	2	3	2	3	4	3	1	2	2	2	3	3	3	4
19	2	3	4	2	3	1	1	1	5	2	1	2	1	2	3	1	1	4	2	2	4	4	4
20	2	1	1	2	3	4	4	1	5	1	2	5	2	5	5	5	2	5	1	5	5	5	5

3. البدء في التحليل : Analyze - Data Reduction - Factor



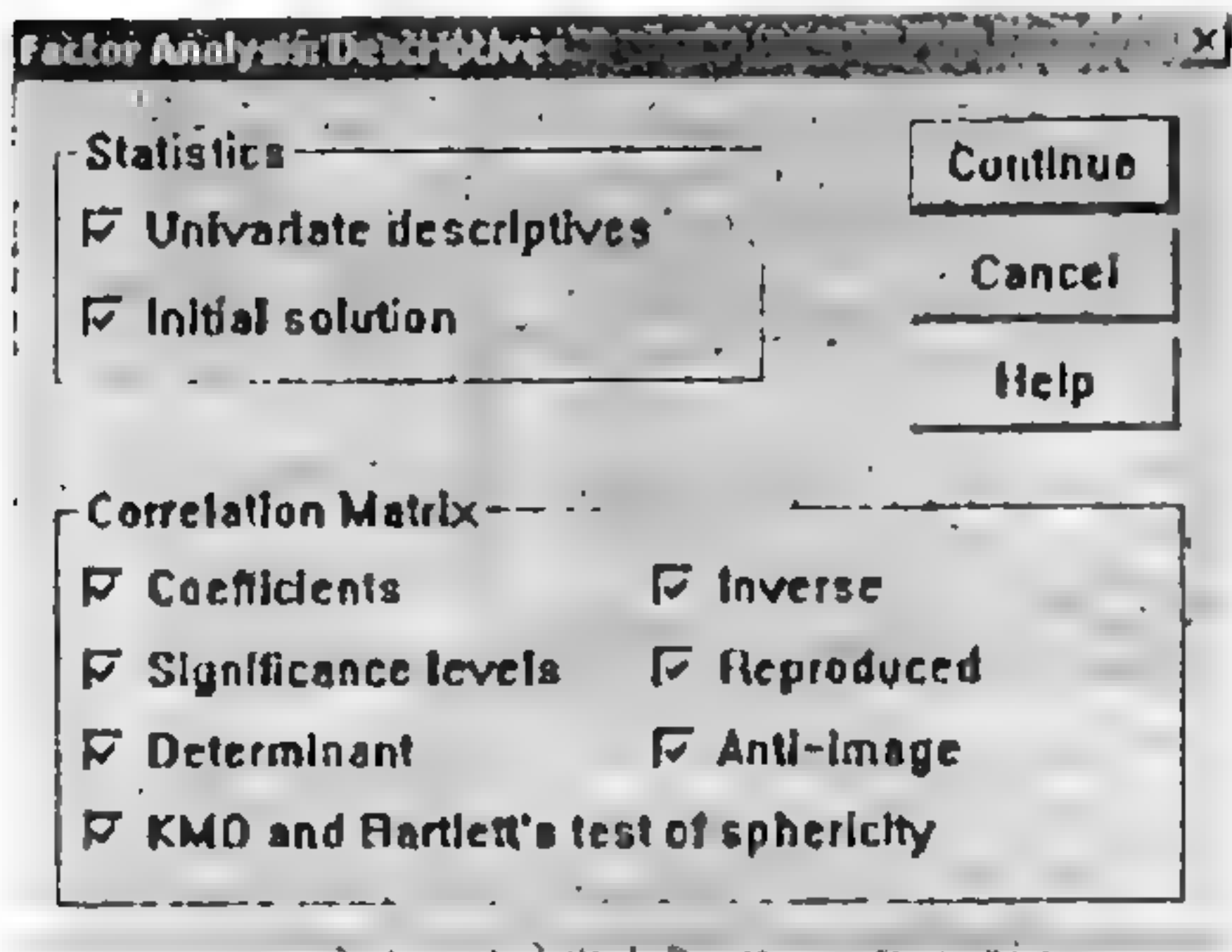
4. اختيار المتغيرات :

نختار المتغيرات ونضعها في قائمة Variables



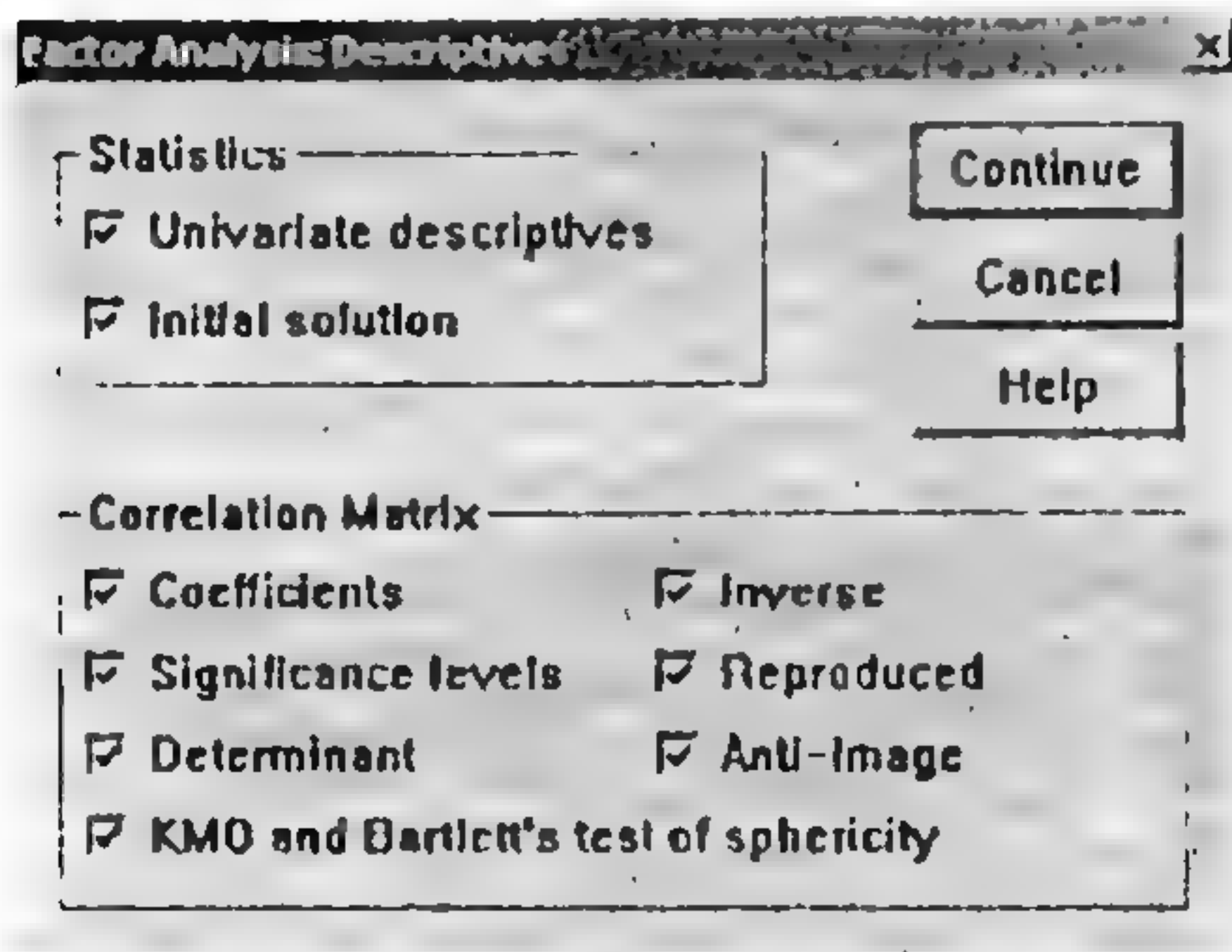
5. إكمال محددات التحليل :

نضغط على زر Descriptives - ثم نحدد Initial solution



- ثم نضغط زر Continue

نضغط على زر Extraction - ثم نحدد:



– ثم نضغط زر Continue

نضغط على زر Rotation... – ثم نحدد:

Factor Analysis: Rotation

Method

☐ None ☐ Quartimax

☒ Varimax ☐ Equamax

☐ Direct Oblimin ☐ Promax

Delta: Kappa:

Display

☒ Rotated solution ☒ Loading plot(s)

Maximum iterations for Convergence:

Continue Cancel Help

– ثم نضغط زر Continue

نضغط على زر Scores... – ثم نحدد:

Factor Analysis: Factor Scores

☒ Save as variables

Method

☐ Regression

☐ Bartlett

☒ Anderson-Rubin

☒ Display factor score coefficient matrix

Continue Cancel Help

– ثم نضغط زر Continue

نضغط على زر Options... – ثم نحدد:

Factor Analysis: Options

Missing Values

☐ Exclude cases listwise

☒ Exclude cases pairwise

☐ Replace with mean

Coefficient Display Format

☒ Sorted by size

☒ Suppress absolute values less than:

Continue Cancel Help

— ثم نضغط زر Continue

6. مخرجات التحليل:

الإحصاءات الوصفية: الوسط الحسابي، والانحراف المعياري للأسئلة.

Descriptive Statistics

	Mean	Std. Deviation	Analysis N	Missing N
Q01	2.37	.828	2571	0
Q02	1.62	.851	2571	0
Q03	2.59	1.075	2571	0
Q04	2.79	.949	2571	0
Q05	2.72	.965	2571	0
Q06	2.23	1.122	2571	0
Q07	2.92	1.102	2571	0
Q08	2.24	.873	2571	0
Q09	2.85	1.263	2571	0
Q10	2.28	.877	2571	0
Q11	2.26	.881	2571	0
Q12	3.16	.916	2571	0
Q13	2.45	.949	2571	0
Q14	2.88	.999	2571	0
Q15	2.77	1.009	2571	0
Q16	2.88	.916	2571	0
Q17	2.47	.884	2571	0
Q18	2.57	1.053	2571	0
Q19	2.29	1.101	2571	0
Q20	3.62	1.036	2571	0
Q21	3.17	.985	2571	0
Q22	2.89	1.041	2571	0
Q23	3.43	1.044	2571	0

مصفوفة الارتسباط بين كل أزواج الأسئلة وتستخدم لفحص العلاقة بين الأسئلة، أنه عندما يوجد ارتباط كبير نسبياً بين علامات اختبارين فمعنى ذلك أن هناك تبايناً مشتركاً بين الاختبارين، أي أنهما يقيسان شيئاً مشتركاً، ونلاحظ أن ارتباطاً بالاختبارات الأخرى متدنية بشكل ملحوظ.

إذن يمكن أن نخلص إلى أن الفكرة الأساسية للنموذج العاملي هي افتراض إمكانية تجميع المتغيرات بناءً على معاملات الارتباط بينها، هذا يعني أن جميع المتغيرات الموجودة في مجموعة معينة مرتبطة مع بعضها ارتباطاً قوياً، ولكن ارتباطها بمتغيرات المجموعات الأخرى ارتباطاً

ضعيفا ، ومن الممكن أن نتصور أن كل مجموعة من المتغيرات تمثل عاملا واحدا وهو المسئول عن الارتباط المشاهد بينها .

Correlation Matrix

		C1	C2	C3	C4	C5	C6	C7
Grand Total	C1	1.000	-.022	-.007	-.002	-.002	-.007	-.009
	C2	-.022	1.000	-.003	-.002	-.002	-.004	-.009
	C3	-.007	-.003	1.000	-.001	-.001	-.002	-.002
	C4	-.002	-.002	-.001	1.000	-.001	-.001	-.001
	C5	-.002	-.002	-.001	-.001	1.000	-.001	-.001
	C6	-.007	-.004	-.002	-.001	-.001	1.000	-.001
	C7	-.009	-.009	-.002	-.001	-.001	-.001	1.000
	C8	-.001	-.001	-.001	-.001	-.001	-.001	-.001
	C9	-.001	-.001	-.001	-.001	-.001	-.001	-.001
	C10	-.001	-.001	-.001	-.001	-.001	-.001	-.001
	C11	-.001	-.001	-.001	-.001	-.001	-.001	-.001
	C12	-.001	-.001	-.001	-.001	-.001	-.001	-.001
	C13	-.001	-.001	-.001	-.001	-.001	-.001	-.001
	C14	-.001	-.001	-.001	-.001	-.001	-.001	-.001
	C15	-.001	-.001	-.001	-.001	-.001	-.001	-.001
	C16	-.001	-.001	-.001	-.001	-.001	-.001	-.001
	C17	-.001	-.001	-.001	-.001	-.001	-.001	-.001
	C18	-.001	-.001	-.001	-.001	-.001	-.001	-.001
	C19	-.001	-.001	-.001	-.001	-.001	-.001	-.001
	C20	-.001	-.001	-.001	-.001	-.001	-.001	-.001
	C21	-.001	-.001	-.001	-.001	-.001	-.001	-.001
	C22	-.001	-.001	-.001	-.001	-.001	-.001	-.001
	C23	-.001	-.001	-.001	-.001	-.001	-.001	-.001
Sig. Partial	C1		1.000	1.000	1.000	1.000	1.000	1.000
	C2			1.000	1.000	1.000	1.000	1.000
	C3				1.000	1.000	1.000	1.000
	C4					1.000	1.000	1.000
	C5						1.000	1.000
	C6							1.000
	C7							
	C8							
	C9							
	C10							
	C11							
	C12							
	C13							
	C14							
	C15							
	C16							
	C17							
	C18							
	C19							
	C20							
	C21							
	C22							
	C23							

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.930
Bartlett's Test of Sphericity	Approx. Chi-Square	19334.492
	df	253
	Sig.	.000

Communalities

	Initial	Extraction
Q01	1.000	.435
Q02	1.000	.414
Q03	1.000	.530
Q04	1.000	.469
Q05	1.000	.343
Q06	1.000	.654
Q07	1.000	.545
Q08	1.000	.739
Q09	1.000	.484
Q10	1.000	.335
Q11	1.000	.690
Q12	1.000	.613
Q13	1.000	.536
Q14	1.000	.488
Q15	1.000	.378
Q16	1.000	.487
Q17	1.000	.683
Q18	1.000	.597
Q19	1.000	.343
Q20	1.000	.484
Q21	1.000	.650
Q22	1.000	.464
Q23	1.000	.412

Extraction Method: Principal Component Analysis.

Total Variance Explained

Component	Total			Extraction			Total		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	7.250	29.804	29.804	7.250	29.804	29.804	7.250	29.804	29.804
2	7.139	29.350	59.154	7.139	29.350	59.154	7.139	29.350	59.154
3	6.725	27.701	86.855	6.725	27.701	86.855	6.725	27.701	86.855
4	6.306	25.624	112.479	6.306	25.624	112.479	6.306	25.624	112.479
5	4.266	17.465	129.944	4.266	17.465	129.944	4.266	17.465	129.944
6	3.929	16.118	146.062	3.929	16.118	146.062	3.929	16.118	146.062
7	3.250	13.400	159.462	3.250	13.400	159.462	3.250	13.400	159.462
8	3.106	12.825	172.287	3.106	12.825	172.287	3.106	12.825	172.287
9	3.004	12.416	184.703	3.004	12.416	184.703	3.004	12.416	184.703
10	2.906	12.026	196.729	2.906	12.026	196.729	2.906	12.026	196.729
11	2.912	12.037	208.766	2.912	12.037	208.766	2.912	12.037	208.766
12	2.811	11.647	220.413	2.811	11.647	220.413	2.811	11.647	220.413
13	2.801	11.621	232.034	2.801	11.621	232.034	2.801	11.621	232.034
14	2.312	9.649	241.683	2.312	9.649	241.683	2.312	9.649	241.683
15	2.350	9.725	251.408	2.350	9.725	251.408	2.350	9.725	251.408
16	2.295	9.581	260.989	2.295	9.581	260.989	2.295	9.581	260.989
17	2.210	9.241	270.230	2.210	9.241	270.230	2.210	9.241	270.230
18	2.200	9.216	279.446	2.200	9.216	279.446	2.200	9.216	279.446
19	1.906	7.943	287.389	1.906	7.943	287.389	1.906	7.943	287.389
20	1.850	7.750	295.139	1.850	7.750	295.139	1.850	7.750	295.139
21	1.850	7.750	302.889	1.850	7.750	302.889	1.850	7.750	302.889
22	1.406	5.863	308.752	1.406	5.863	308.752	1.406	5.863	308.752
23	1.406	5.863	314.615	1.406	5.863	314.615	1.406	5.863	314.615

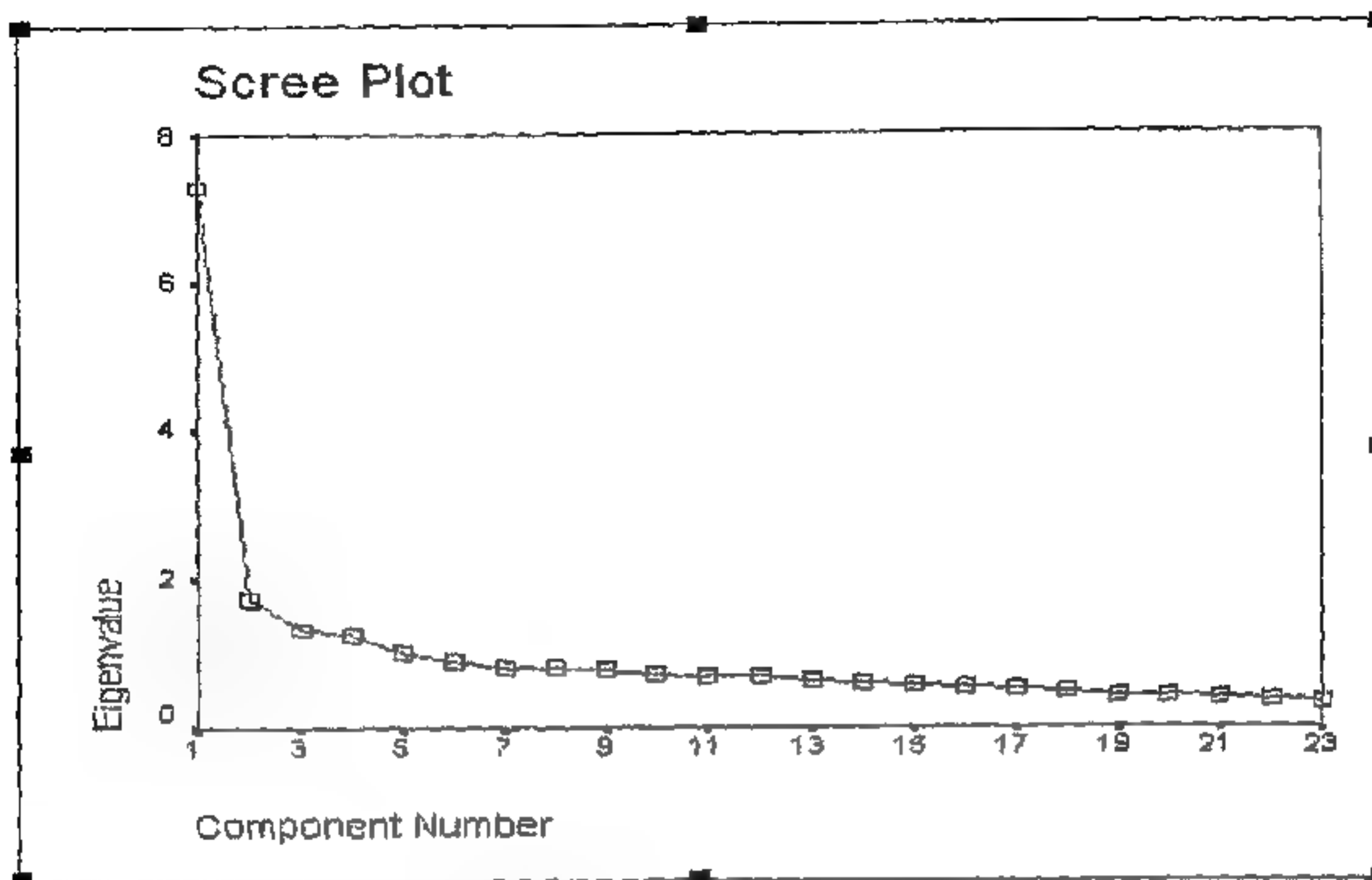
Extraction Method: Principal Component Analysis.

جدول رقم (2)

العوامل المكونة للأداة وجذورها الكامنة ونسبة التباين العاملة

رقم العامل	الجذر الكامن	نسبة التباين العاملة	نسبة التباين التراكمية
1	7.290	31.696	31.696
2	1.739	7.560	39.256
3	1.317	5.725	44.981
4	1.227	5.336	50.317
22	0.364	1.583	98.562
23	0.333	1.448	100.00

وحيث تم اجراء التحليل العاملي كأحد إجراءات تحقيق صدق الأداة لذا يتضح من النتائج الواردة في جدول رقم (2) أن الأداة تتكون من ثلاث وعشرون عاملاً تشبع عليها بنود الأداة بقيم تفوق (0.30) حسب محك جيلفورد أما الجذور الكامنة للعوامل فتتراوح بين 7.290 و 0.333 بينما كانت نسبة التباين العاملة بين 31.696 و 1.448 وهذه قيم مقبولة حسب محك كيزر. وهذه العوامل مجتمعة تفسر ما نسبته (50.317%) من الظاهرة وهذه النسبة عالية خاصة إذا علمنا أن نسبة (10%) تعتبر نسبة مقبولة.



Component Matrix^a

	Component			
	1	2	3	4
Q18	.701			
Q07	.686			
Q16	.679			
Q13	.673			
Q12	.669			
Q21	.668			
Q14	.668			
Q11	.652			-.400
Q17	.649			
Q04	.634			
Q03	-.628			
Q15	.593			
Q01	.585			
Q06	.585			
Q08	.546	.401		-.417
Q10	.437			
Q20	.436		-.404	
Q19	-.427			
Q09		.827		
Q02		.648		
Q22		.485		
Q05	.662		.571	
Q23				.527

Extraction Method: Principal Component Analysis.

a. 4 components extracted.

Rotated Component Matrix^a

	Component			
	1	2	3	4
Q06	.800			
Q18	.684			
Q13	.647			
Q07	.638			
Q14	.579			
Q10	.550			
Q15	.469			
Q20		.677		
Q21		.661		
Q03		.667		
Q12	.473	.523		
Q04		.516		
Q16		.514		
Q01		.486		
Q05		.428		
Q08			.833	
Q17			.747	
Q11			.747	
Q09				.648
Q22				.645
Q23				.586
Q02				.543
Q19				.428

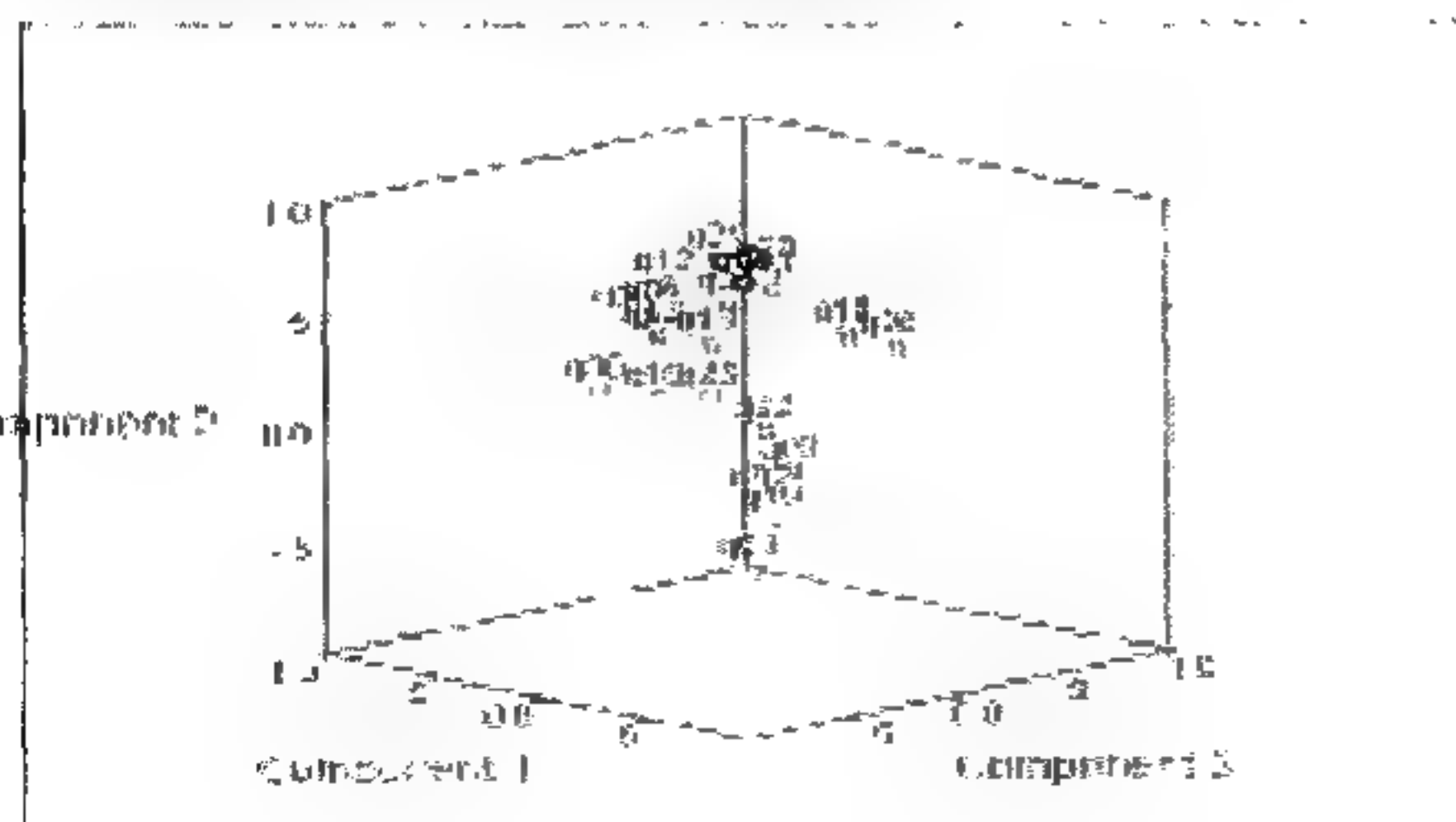
Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

Component Transformation Matrix

Component	1	2	3	4
1	.635	.585	.443	-.242
2	.137	-.167	.488	.846
3	.758	-.513	-.403	.008
4	.087	.605	-.835	.476

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization

Component Plot in Rotated Space



8-11 تمارين Exercise

إذا كان لديك البيانات التالية جد العوامل التي تكونها البيانات.

	idno	sex	spec	q1	q2	q3	q4	q5	q6	q7	q8	q9	q10	q11	q12	q13	q14	q15
1	1	1	3	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1
2	2	1	3	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
3	3	1	3	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1
4	4	1	3	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1
5	5	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1
6	6	1	1	1	1	0	1	1	0	1	1	1	1	1	0	1	1	1
7	7	1	1	1	1	0	0	1	0	1	1	1	0	1	0	1	1	1
8	8	1	1	1	0	0	0	1	0	1	1	1	0	1	1	1	1	1
9	9	1	1	1	0	0	1	1	0	1	1	1	1	0	1	1	1	1
10	10	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	1
11	11	2	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
12	12	2	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
13	13	2	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1
14	14	2	1	0	1	0	0	0	0	0	1	1	0	0	0	1	1	1
15	15	2	1	0	0	1	0	0	0	1	1	1	1	1	1	1	1	1
16	16	2	2	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1
17	17	2	2	1	1	0	1	1	0	0	1	1	1	1	1	1	1	1
18	18	2	2	0	0	0	1	0	1	1	1	1	1	1	0	0	1	0
19	19	2	2	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1
20	20	2	2	1	1	0	0	1	0	1	1	1	1	0	0	0	1	1
21	21	2	2	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1
22	22	2	2	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1
23	23	2	2	1	0	0	0	0	0	1	1	1	0	1	1	1	1	1
24	24	2	2	0	0	0	1	0	1	0	0	1	0	0	1	0	1	0
25	25	2	3	1	0	0	0	1	0	0	1	1	0	0	0	1	0	1
26	26	2	3	1	0	0	1	1	0	1	1	1	1	0	1	1	1	1
27	27	2	3	1	1	0	1	1	0	1	0	1	1	1	0	1	1	1
28	28	2	3	1	1	0	1	0	0	1	0	1	1	1	0	1	1	1
29	29	2	3	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1
30	30	2	3	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1

الفصل الثاني عشر

الإحصاءات الالاعلمية

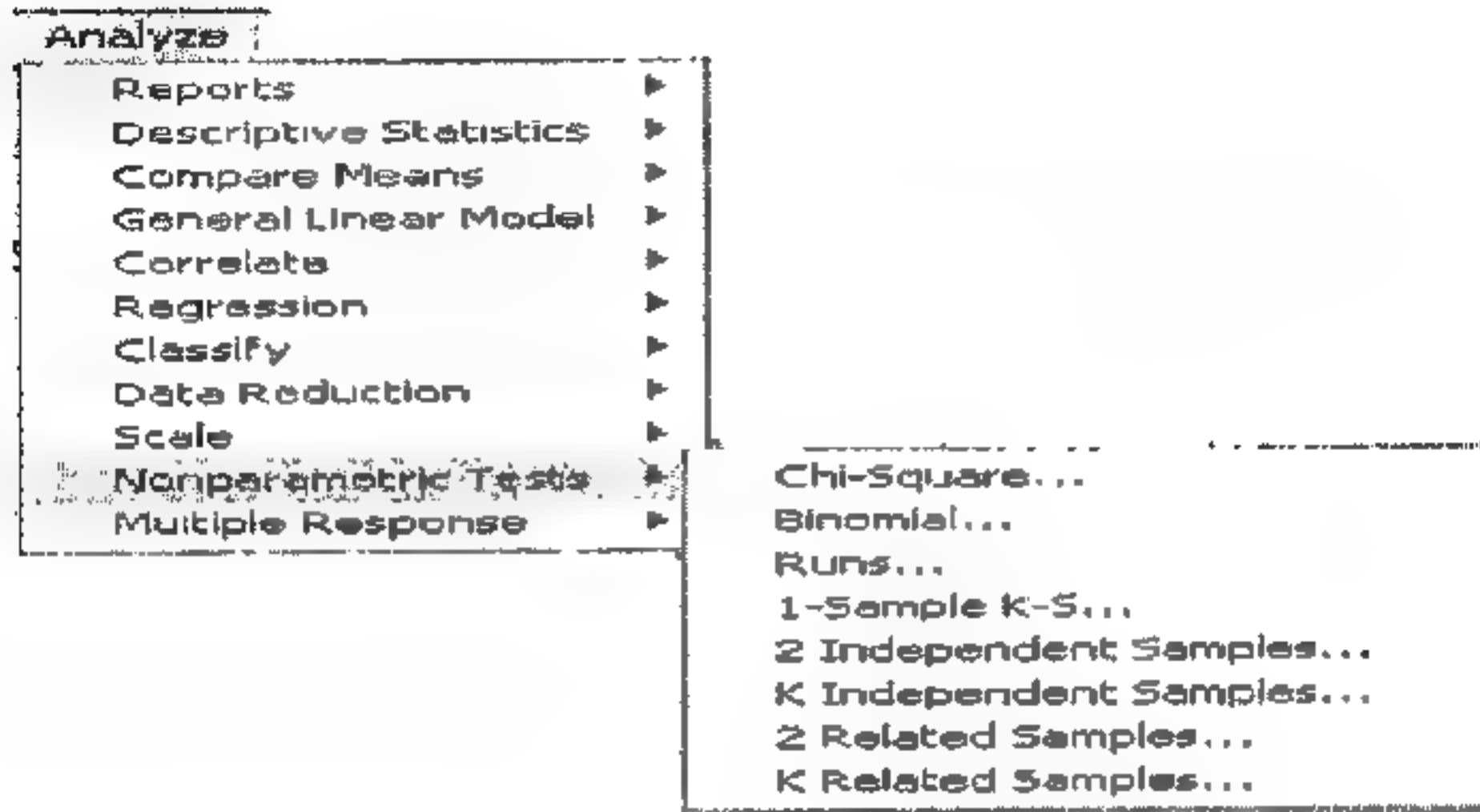
NONPARAMETRIC STATISTICS

- 1-12 مقدمة Introduction
- 2-12 الطرق الالاعلمية (عينة واحدة)
Nonparametric Methods (Single Sample)
- 3-12 الطرق الالاعلمية (عينتين مستقلتين)
Nonparametric Methods (Two Independent Samples)
- 4-12 الطرق الالاعلمية (عينتين مرتبطتين)
Nonparametric Methods (Two Related Samples)
- 5-12 الطرق الالاعلمية (ثلاثة عينات مستقلة أو أكثر)
Nonparametric Methods (3 or more Independent Samples)
- 6-12 الطرق الالاعلمية (ثلاثة عينات مرتبطة أو أكثر)
Nonparametric Methods (3 or more Related Samples)
- 7-12 استخدام برمجية SPSS في حل المسائل.
- 8-12 تمارين Exercise

الفصل الثاني عشر

الإحصاءات الالاعلمية

NONPARAMETRIC STATISTICS



1-12 مقدمة Introduction

إذا كانت المتغيرات التابعة مقاسة بمقياس اسمي أو رتبي أو عندما لا نتمكن من الإيفاء بافتراضات الاختبارات المعلمية، فإنه يمكن استخدام الاختبارات الالاعلمية Nonparametric tests لأن مثل هذه الاختبارات لا تتطلب أية افتراضات حول المجتمعات الإحصائية مثل التوزيع الطبيعي وتجانس التباين واختيار العينة من المجتمع عشوائياً.

والاختبارات الالاعلمية بصورة عامة أكثر قوة من الاختبارات المعلمية، إذ أن الإحصاءات المعلمية تميل أكثر من الإحصاءات الالاعلمية لرفض الفرضية الصفرية، وأن الاختبارات الالاعلمية أسهل في طريقة إجراؤها.

الاختبارات الالاعلمية هي أساليب متحررة من التوزيع Distribution Free

Methods وأساليب متحررة من الافتراضات Assumption Free Methods

الإحصاءات الالاعلمية قائمة على افتراضات ضعيفة، بينما الإحصاءات المعلمية قائمة على افتراضات قوية وأساسية، فإذا تحققت الافتراضات فإن الإحصاءات المعلمية أكثر فاعلية من

الإحصاءات الالامعلمية، وإذا لم تتحقق الافتراضات فإن الإحصاءات الالامعلمية أكثر فاعلية من الإحصاءات العلمية.

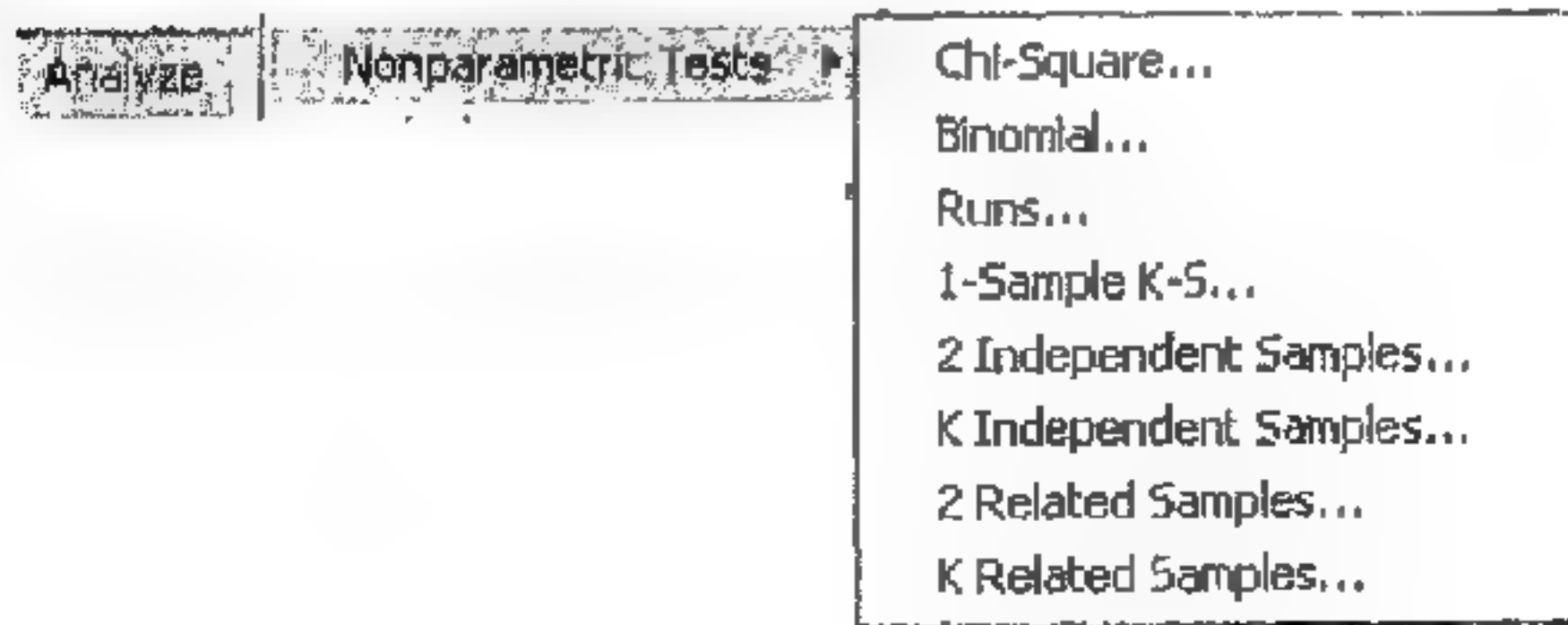
الاسلوب الالامعلمي باختصار هو اسلوب احصائي يمتلك خصائص معينة مرغوبة تحت افتراضات ضعيفة نسبياً تخص المجتمعات المعنية التي منها تم الحصول على البيانات.

* متى نستخدم الإحصاءات الالامعلمية

1. الفرضية المراد اختبارها لا تتعلق بمعالم المجتمع.
2. البيانات المتوفرة مقاسة وفق تدرج أضعف من متطلب الاساليب العلمية.
3. الافتراضات الخاصة بالاستخدام الملائم للاساليب العلمية غير متحقق (انتهاك واضح لافتراضات الاساليب العلمية).
4. الحصول على نتائج بسرعة ومهولة وتكلفة أقل.

2-12 الطرق الالامعلمية (عينة واحدة) Nonparametric Methods (Single Sample)

1. اختبار الإشارة Sign test
2. اختبار ويلكسون للرنب ذات الإشارة Wilcoxon signed ranks test
3. اختبار السلاسل Runs test for randomness
4. اختبار كوكس ستيوارت للاتجاه Cox-Stuart test for trend
5. اختبار ذات الحدين Binomial test
6. اختبار كولمقروف سميرنوف Kolmogrov-Smirnov Test



12-2-1 اختبار الإشارة (عينة واحدة) Sign test (One sample)

* الافتراضات Assumptions

1. العينة مختارة عشوائياً من مجتمع وسيطه غير معلوم.
2. المتغير المقصود مقياس على مقياس رتبي على الأقل.
3. توزيع الاختبار الاحصائي هو توزيع ذات الحدين $B(n, 1/2)$

* الفرضية المطلوب اختبارها Hypothesis

الفرضية الصفرية: الوسيط = مقدار ثابت $H_0: \theta = \theta_0$

الفرضية البديلة: الوسيط \neq مقدار ثابت $H_1: \theta \neq \theta_0$

* الاختبار الاحصائي Test Statistic

الاختبار الاحصائي هو العدد الأقل للإشارات الموجبة أو السالبة

$$T = r(\text{smallest no of } + \text{ or } - \text{ signs})$$

* اتخاذ القرار Decision Rule

نرفض الفرضية الصفرية اذا كانت القيمة المحسوبة $P(T \leq r)$ أقل من القيمة الحرجة $\alpha/2$

نقبل الفرضية الصفرية اذا كانت القيمة المحسوبة $P(T \leq r)$ أكبر من القيمة الحرجة $\alpha/2$

مثال 12-1: اختبرت عينة عشوائية تتكون من 20 كتاباً وكانت اعداد الصفحات هذه

الكتب كما يلي:

153, 166, 181, 192, 244, 248, 258, 264, 266, 305, 305, 312, 330, 340,
356, 361, 395, 427, 433, 467

استخدم هذه البيانات في اختبار الفرضية الصفرية، عند مستوى دلالة $(\alpha=0.05)$.

$$H_0: \theta = 250$$

$$H_1: \theta \neq 250$$

3-12 الطرق الالاعلمية (عينتين مستقلتين) Nonparametric Methods (Two Independent Samples)

1. اختبار توكي السريع (TQT) Tukey's Quick Test
2. اختبار الوسيط Median Test
3. اختبار مان-وتني Mann-Whitney Test
4. اختبار تساوي التباينات - اختبار انصاري-برادلي Test for equality of variance (Ansari-Bradley Test)
5. اختبار تساوي التباينات - اختبار موسى Test for equality of variance (Moses Test)
6. اختبار تساوي التباينات - اختبار مربعات الرتب Test for equality of variance (The Sequare Rank Test)
7. اختبار التوزيعات المتماثلة - اختبار سميرنوف Test fir identical distribution (Smirnov Test)

* اختبارات تساوي الوسيطات لعينتين مستقلتين:

1. اختبار توكي السريع (TQT) Tukey's Quick Test
 $H_0: M_1 = M_2$
2. اختبار الوسيط Median Test
 $H_0: M_1 = M_2$
3. اختبار مان-وتني Mann-Whitney Test
 $H_0: M_1 = M_2$

* اختبارات تساوي التباينات لعينتين مستقلتين

4. اختبار تساوي التباينات - اختبار انصاري-برادلي
 $H_0: \sigma_1^2 = \sigma_2^2$
5. اختبار تساوي التباينات - اختبار موسى
 $H_0: \sigma_1^2 = \sigma_2^2$
6. اختبار تساوي التباينات - اختبار مربعات الرتب
 $H_0: \sigma_1^2 = \sigma_2^2$

* اختبارات التماثل لعينتين مستقلتين

التماثل: التساوي في مقاييس التوزعة المركزية، التشتت، التماثل في الشكل، وإحلال أي منهم يجعل التوزيعات غير متماثلة.

7. اختبار التوزيعات المتماثلة - اختبار سميرنوف
 $F_1(x) = F_2(x)$

اختبارات تساوي الوسيطات لعينتين مستقلتين - اختبار مان-وتني Mann-Whitney
يستخدم لأغراض مقارنة وسيطين لمجتمعين بالاعتماد على مشاهدات مستقلة تختار عشوائياً من المجتمعين.
يستخدم للمقارنة بين عينتين مستقلتين عندما تكون البيانات عديدة بطبيعتها، وهو البديل الالامعلمي لاختبار T المعلمي للبيانات المستقلة.
والفرضية الصفرية له: لا يوجد اختلاف جوهري في علامات الأفراد في المجتمعات التي سحبت منها العينتان.

* افتراضات الاختبار

1. تكون البيانات من عينة من المشاهدات العشوائية x_1, x_2, \dots, x_{n1} من المجتمع الأول الذي وسيطه $M1$ غير معروف، ومن عينة أخرى من المشاهدات y_1, y_2, \dots, y_{n2} من المجتمع الثاني الذي وسيطه $M2$ غير معروف.
2. المتغيرات متصلة x, y .
3. مقياس على سلم رتبي على الأقل.

* الفرضية Hypothesis

وسيط المجتمع الأول = وسيط المجتمع الثاني $H_0: M1 = M2$
وسيط المجتمع الأول \neq وسيط المجتمع الثاني $H_0: M1 \neq M2$

* لاختبار الفرضية التي تتعلق بتساوي وسيطين نقوم بالآتي:

1. نشكل العينة الكلية من العينتين.
2. نرتب المشاهدات في العينة الكلية من الأقل إلى الأكبر ونعطيها الرتب التي تبدأ من 1 إلى $n1+n2$ لأكثر قيمة.
3. نجمع الرتب الخاصة بمشاهدات العينة المأخوذة من المجتمع الأول ويرمز للمجموع بالرمز S.
4. نحسب الاختبار الاحصائي T كما يلي:

$$T = S - \frac{n_1(n_1+1)}{2}$$

* القرار:

رفض H_0 إذا كان قيمة الاختبار الإحصائي T أكبر من القيمة الحرجة $W_1 - (\alpha/2)$

إذا كان قيمة الاختبار الإحصائي T أقل من القيمة الحرجة $W(\alpha/2)$

حيث نجد القيمة الحرجة $W(\alpha/2)$ من الجدول A7

$$W_1 - (\alpha/2) = n_1 n_2 - W(\alpha/2) \quad \text{وتكون}$$

* ملاحظة: إذا كانت أي منهما أو كليهما أكبر من 20 نستخدم التقريب للتوزيع

الطبيعي.

$$Z = \frac{T - (n_1 n_2 / 2)}{\sqrt{n_1 n_2 (n_1 + n_2 + 1) / 12}}$$

مثال 12-2: في دراسة حول تقييم الصدق التمييزي لاختبار ماء، تم تقسيم مجموعة من

الأفراد إلى مجموعتين متمايزتين في السمة التي يقيسها الاختبار، حيث يمتلك أفراد المجموعة الأولى

والبالغ عددهم 17 فرداً السمة بدرجة عالية، بينما يمتلك أفراد المجموعة الثانية والبالغ عددهم

10 أفراد السمة بدرجة منخفضة، إذا طبق الاختبار على أفراد المجموعتين وكانت العلامات

لأفراد المجموعتين كما في الجدول التالي:

هل يمتلك الاختبار قدرة تمييزية عند مستوى دلالة $(\alpha = 0.05)$.

* الحل:

MARK-G1	MARK-G2	Sort-G1	Sort-G2	Rank-G1	Rank-G2
11.80	8.60	11.80	8.60	27.00	15.00
11.70	5.80	11.70	5.80	26.00	13.00
9.50	5.40	9.50	5.40	25.00	12.00
9.40	5.10	9.40	5.10	24.00	11.00
8.70	5.00	8.70	5.00	23.00	9.50
8.20	4.30	8.20	4.30	22.00	8.00
7.70	3.90	7.70	3.90	21.00	5.00
7.40	3.30	7.40	3.30	19.50	4.00
7.40	2.40	7.40	2.40	19.50	3.00
7.10	1.70	7.10	1.70	18.00	1.00
6.90		6.90		17.00	
		6.80		16.00	
			6.60		
6.30		6.30		14.00	
			5.80		
			5.40		
			5.10		
			5.00		
5.00		5.00		9.5	
			4.30		
4.20		4.20		7	
4.10		4.10		6	
			2.4		
2.20		2.20		2	
			1.7		
count	17	10		296.50	61.50
sum					

بمجموع الرتب للعينه الأولى S1

$$S1=27+26+25+24+23+22+21+19.5+18+17+16+14+9.5+7+6+2=296.5$$

$$T = S - \frac{n_1(n_1+1)}{2} = 296.5 - (18*17)/2 = 296.5-153 = 143.5$$

$$46 = W(\alpha/2) \text{ نجد أن } n1=10, n2=17, P=0.025$$

$$W_1 - (\alpha/2) = n1n2 - W(\alpha/2) = 18*17 - 46 = 260$$

* القرار:

رفض H_0 اذا كان قيمة الاختبار الاحصائي T اكبر من القيمة الحرجة $W_1 - (\alpha/2)$

اذا كان قيمة الاختبار الاحصائي T أقل من القيمة الحرجة $W(\alpha/2)$

اذا نرفض الفرضية الصفرية H_0 ، وهذا يعني أن هناك قذوة تمييزية للاختبار.

بمجموع الرتب للعينه الثانية S2

$$S2=12+13+15+11+9.5+8+5+4+3+1=81.5$$

$$T = S - \frac{n_2(n_2+1)}{2} = 81.5 - (10*11)/2 = 26.5$$

$$46 = W(\alpha/2) \text{ نجد أن } n1=10, n2=17, P=0.025$$

$$W_1 - (\alpha/2) = n1n2 - W(\alpha/2) = 10*17 - 46 = 124$$

القرار: بما أن قيمة الاحصائي $T=26.5$ وهي خارج حدود الفترة (46 - 124) اذا

نرفض الفرضية الصفرية H_0

التقريب:

$$Z = \frac{T - (n1n2/2)}{\sqrt{n1n2(n1+n2+1)/12}} = \frac{26.5 - 85}{\sqrt{10*17(10+17+1)/12}} = \frac{-58.5}{19.916} = -2.937$$

القرار: بما أن القيمة المحسوبة $Z = -2.937$ أقل من القيمة الحرجة نرفض الفرضية

الصفرية

12-4 الطرق الالاعلمية (عينتين مرتبطين) Nonparametric Methods (Two Related Samples)

1. اختبار الاشارة لعينتين مرتبطين Sign Test
2. اختبار ويلكسون لاشارة رتب الفروق المطلقة Wilcoxon signed ranks test
3. اختبار مكنمار لعينتين مرتبطين Mc Nemar's test
4. اختبار مربع كاي χ^2 للاستقلالية Chi-square test of independence
5. اختبار مربع كاي χ^2 للتجانس Chi-square test of homogeneity

* اختبارات الوسيطات لعينتين مترابطتين:

1. اختبار الاشارة لعينتين مرتبطين $H_0: M_1 = M_2$
2. اختبار ويلكسون لاشارة رتب الفروق المطلقة $H_0: M_1 = M_2$

* اختبارات النسب لعينتين مترابطتين:

3. اختبار مكنمار لعينتين مرتبطين $H_0: P_1 = P_2$

* اختبارات الاستقلال لعينتين مترابطتين:

4. اختبار مربع كاي χ^2 للاستقلالية

* اختبارات التجانس لعينتين مترابطتين:

5. اختبار مربع كاي χ^2 للتجانس

* اختبارات الوسيطات لعينتين مترابطتين:

* اختبار ويلكسون لاشارة رتب الفروق المطلقة Wilcoxon signed ranks test

تستخدم في التصاميم التجريبية ذات الاختبارين القبلي والبعدي، وهو البديل الالاعلمي لاختبار T ببيانات المترابطة، ومن أهم ميزات هذا الاختبار انه يختبر اتجاه الفرق بين أزواج المشاهدات وحجم هذا الفرق النسبي أيضاً، ويجب أن تكون المشاهدات رقمية، ولا يمكن استخدامه إذا كانت تصنيفية اسمية.

إذا كانت الفروق تفضل إحدى المجموعتين تكون هذه المجموعة هي الأفضل بدلالة إحصائية.

* الافتراضات

1. البيانات تلخص بمجموعة من الأزواج العشوائية.
2. يمثل الفرق بين العنصر الأول والثاني في الأزواج المرتبة متغيراً متصلاً.
3. توزيع الفروق متماثل حول الوسيط للفرق.
4. الفروق مستقلة من خلال الاختيار العشوائي للأفراد.
5. الفروق مقاسة على الأقل في المستوى الفئوي.

* الفرضية

$$H_0: M_D = 0, M_1 = M_2$$

$$H_1: M_D \neq 0, M_1 \neq M_2$$

* لأختبار الفرضية نعمل الآتي:

1. احصل على الفروق $X_i - Y_i$
2. اعطي رتباً للفروق المطلقة R
3. اعطي الرتب اشارات الفروق التي انتجتها
4. احسب مجموع الرتب ذات الإشارة الموجبة $T+$ ومجموع الرتب السالبة $T-$

$$T = K = \text{smaller } (T+ \text{ or } T-)$$

* الاختبار الاحصائي

* القرار

ارفض الفرضية الصفرية إذا كانت قيمة الاختبار الاحصائي $T \geq$ القيمة الحرجة C التي يمكن الحصول عليها من جدول A.3 ، أو ارفض الفرضية الصفرية إذا كان $P(T=K|n) \leq (\alpha/2)$ وفيما عدا ذلك نقبل الفرضية الصفرية.

* ملاحظة

عندما يكون من غير الممكن استخدام جدول A.3 بسبب أن $n < 30$ نستخدم القيمة التقريبية من التوزيع الطبيعي المعياري للاختبار والتي تحسب من المعادلة التالية:

$$Z = \frac{T - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$$

n: عدد الأزواج.

ويتم مقارنة القيمة الناتجة مع القيمة الحرجة 1.96

* مثال 12-3:

درس العالمان لاتاني وكابل أثر اجتماع (التقاء) مجموعة من الفئران على معدل دقات القلب لها وقد رصدوا هذه المعدلات لعشرة فئران بيضاء عندما كان كل منها لوحده وعندما كان مع فار آخر. ويبين الجدول التالي تلك المعدلات:

R	R_R	Di	Di	Rank Di	رتب مع الإشارة
402.00	437.00	-35.00	35	7	-7
409.00	470.00	-61.00	61	9	-9
415.00	408.00	7.00	7	2	2
418.00	448.00	-30.00	30	5	-5
426.00	454.00	-28.00	28	4	-4
450.00	476.00	-26.00	26	3	-3
456.00	535.00	-79.00	79	10	-10
462.00	461.00	1.00	1	1	1
462.00	494.00	-32.00	32	6	-6
463.00	523.00	-60.00	60	8	-8

$$H_0: M_D = 0, M_1 = M_2$$

الفرضية الصفرية: لا يوجد أثر لوضع الفئران مع بعضها على دقات القلب

$$H_1: M_D \neq 0, M_1 \neq M_2$$

الفرضية البديلة: يوجد أثر لوضع الفئران مع بعضها على دقات القلب

* الاختبار الاحصائي

$$T = K = \text{smaller } (T+ \text{ or } T-) = \sum r+ = 1+2 = 3$$

$$P(T=3 | n=10) = 0.0049$$

* القرار

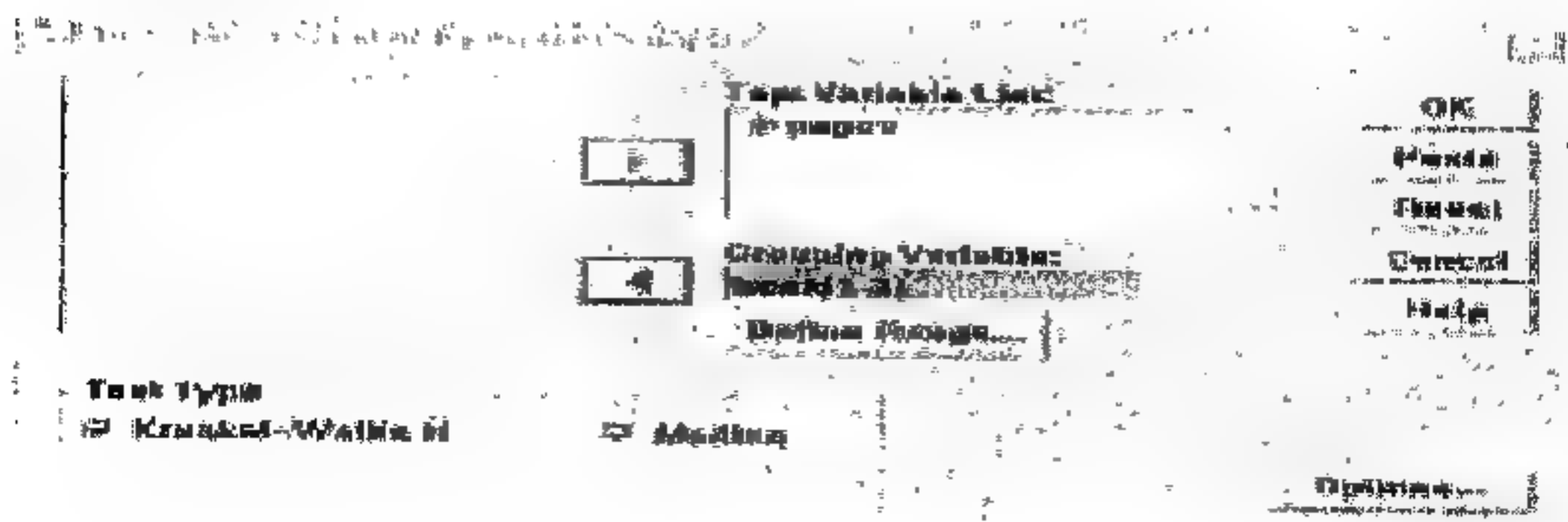
ارفض الفرضية الصفرية اذا كانت قيمة الاختبار الاحصائي $T \geq$ القيمة الحرجة C التي يمكن الحصول عليها من جدول A.3

أو ارفض الفرضية الصفرية اذا كان $P(T=K|n) \leq (\alpha/2)$ كما أن $P(T=3 | n=10) = 0.0049$ أقل من $(\alpha/2) = 0.025$ لذلك نرفض الفرضية الصفرية، ونستنتج أن هناك أثر لوضع الفئران معاً على دقات قلبها.

5-12 الطرق الالامعلمية (ثلاثة عينات مستقلة أو أكثر) Nonparametric Methods(3 or more Independent Samples)

1. اختبار كروسكال والس Kruskal-Wallis test
2. اختبار العلامات الطبيعية البديلة A normal scores alternative test
3. اختبار جونكير في حالة الفرض البديل المرتب Jonckheere test for ordered alternative
4. اختبار تجانس التباين Test for equality of variance

1. اختبار كروسكال والس $H_0: M_1 = M_2 = M_3$
2. اختبار العلامات الطبيعية البديلة $H_0: M_1 = M_2 = M_3$
3. اختبار جونكير في حالة الفرض البديل المرتب $H_0: M_1 = M_2 = M_3$
4. اختبار تجانس التباين $H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2$



* اختبار كروسكال والس Kruskal-Wallis Test

هذا الاختبار يناظر تحليل التباين الأحادي One-Way ANOVA في الاحصاءات المعلمية، وهو تحليل تباين احادي للرتب، وهو امتداد مباشر لاختبار ويلكسون واختبار مان-وتني.

وهو مصمم لاختبار الفرضية الصفرية التي تنص على عدة عينات مستقلة قد سحبت من نفس المجتمع، لذا فهو مصمم لاختبار دلالة الفرق بين ثلاث مجموعات مستقلة أو أكثر عندما تكون البيانات على المتغير التابع رتبية أو يمكن ترتيبها.

Sample1	Sample2	Sample I	Sample k
X_{11}	X_{12}				
X_{21}	X_{22}				
.					
.					
X_{n11}	X_{n22}				

n_1 : حجم العينة الأولى.

n_2 : حجم العينة الثانية.

n_k : حجم العينة k

S_1 : مجموع رتب العينة الأولى.

S_2 : مجموع رتب العينة الثانية. S_k : مجموع رتب العينة K.

* الفرضية الصفرية:

$$H_0: M_1 = M_2 = M_3$$

$$H_1: M_1 \neq M_2 \neq M_3$$

* الطريقة:

1. لدينا n من القيم نرتبها من أقل قيمة لأكبر قيمة (العينة الكلية).
2. نعطيهما رتب من اقل قيمة رتبة 1 واعلى قيمة الرتبة n ، وتعطى القيم المتساوية نفس الرتبة.
3. تعطى الرتب لمواقع المشاهدات في عيناتها.
4. نجمع الرتب الخاصة بكل عينة S_i وهي مجموع الرتب للعينة i

$$5. \text{ نربع مجاميع الرتب لكل عينة } S_t^2 = \sum \frac{S_i^2}{n_i} = \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2} + \dots + \frac{S_k^2}{n_k}$$

6. في حالة عدم وجود ties

$$S_r^2 = \sum r_{ij}^2 = 1^2 + 2^2 + 3^2 + \dots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

* الاختبار الاحصائي

بوجود عدد قيم متعادلة (رتب متساوية) وهي وجود ties

$$T = \frac{(n-1)(S_r^2 - c)}{S_t^2 - c}$$

$$\text{Where } c = 0.25 * n(n+1)^2$$

عدم وجود عدد قيم متعادلة (رتب متساوية) وهي وجود ties

$$T = \frac{12}{n(n+1)} S_t^2 - 3(n+1)$$

* القرار: اذا كان عدد العينات 3 والحجوم فيها 5 أو أقل نستخدم جدول A.12

المحسوبة > الحرجة نفشل في ان نرفض الفرضية الصفرية.

المحسوبة < الحرجة نرفض الفرضية الصفرية.

* ملاحظة: في الحالة التي لا نستطيع فيها استخدام الجدول A.12 حيث (عدد العينات

أكثر من 3 أو حجم العينات في أي منها أكثر من 5) نتوقع أن تتوزع قيم الاختبار T حسب

توزيع χ^2 بدرجات حرية تساوي k-1 حيث (k عدد العينات). لذلك يمكن استخدام جدول

χ^2 في هذه الحالة لمعرفة القيم الحرجة.

مثال 4-12:

تم حصر عدد الصفحات في ثلاثة مجموعات عشوائية من الكتب (احصاء، رياضيات،

حاسوب) واعتماداً على اعداد الصفحات المسجلة نرغب باختبار الفرضية الصفرية التي تنص

على أن العينات الثلاث من نفس التوزيع، مقابل الفرضية البديلة التي تنص على أن هناك على

الأقل عينة واحدة جاءت من توزيع له وسيط مختلف عن الوسيطات في التوزيعات الأخرى عند

مستوى الدلالة الاحصائية $\alpha = 0.05$ نفذ هذا الاختبار اذا كانت عدد الصفحات كما يلي:

	Stat	Math	Comp	Stat_R	Math_R	Comp_R
	126	93	29	10	6	1
	142	98	39	11	7	2
	156	216	60	12	15	3
	228	249	78	17	20	4
	245	301	82	18	25	5
	246	319	112	19	26	8
	370	731	125	28	34	9
	419	910	170	29	36	13
	433		192	30		14
	454		224	31		16
	478		263	32		21
	503		275	33		22
			276			23
			286			24
			369			27
			756			35
مجموع الرتب				270	169	227
حجم العينة				12	8	16

$$K=3 \quad S_3=227 \quad S_2=169 \quad S_1=270$$

$$N=36 \quad n_3=16 \quad n_2=8 \quad n_1=12$$

H0: الفرضية الصفرية: مجموعات الكتب الثلاثة جاءت من نفس التوزيع

H1: الفرضية البديلة: على الأقل مجموعة واحدة جاءت من توزيع له وسط مختلف

$$S^2_t = \sum \frac{S_i^2}{n_i} = \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2} + \frac{S_3^2}{n_3} = \frac{(227)^2}{16} + \frac{(169)^2}{8} + \frac{(270)^2}{12} = 12865.7$$

$$T = \frac{12}{36(37)} S^2_t - 3(n+1) = \frac{12}{36*37} * 12865.7 - 3(37) = 4.91$$

* القرار

بما أن الحجم لبعض العينات تزيد عن 5 فإن توزيع الاختبار الإحصائي T يقترب من

توزيع χ^2 بدرجات حرية (2 = 3-1) لذلك نقبل في رفض الفرضية الصفرية لأن المحسوبة

(4.91) > الحرجة $\chi^2_{2,0.95} = 5.991$ ونستنتج أن مجموعات الكتب الثلاثة من نفس

التوزيع من حيث عدد صفحات الكتاب، أي لها نفس الوسيطات ولها نفس التشتت ونفس الشكل.

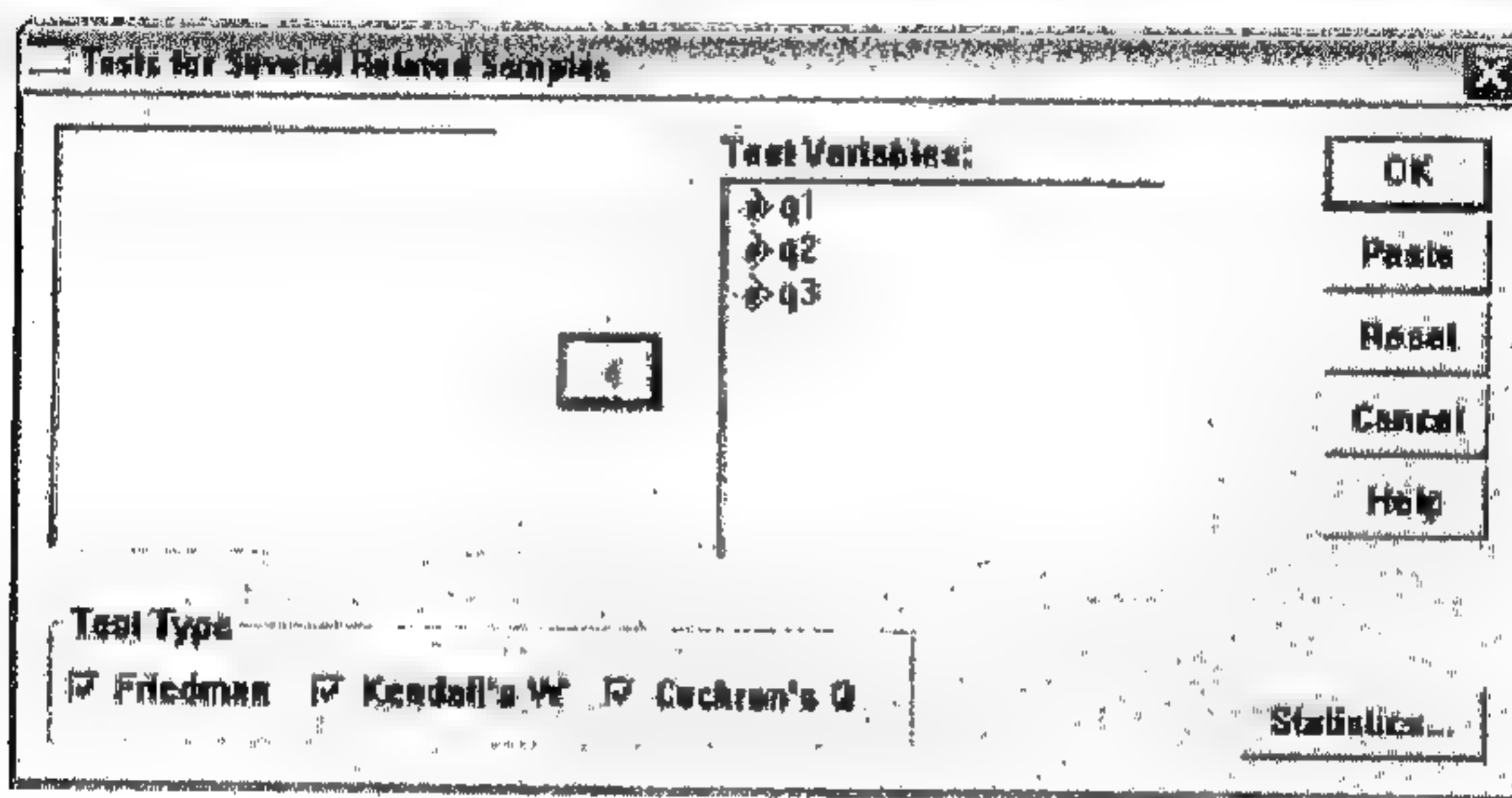
6-12 الطرق اللامعلمية (ثلاثة عينات مرتبطة أو أكثر) Nonparametric Methods (3 or more Related Samples)

$$H_0: M_1 = M_2 = M_3$$

1. اختبار فريدمان Friedman test

$$H_0: M_1 = M_2 = M_3$$

2. اختبار كوكران Cochran's test



* اختبار كوكران Cochran's test

يستخدم هذا الاختبار في حالة البيانات الثنائية (متغير ثنائي) مثل النجاح والفشل، الحدوث وعدم الحدوث، الربح والخسارة، الذكور والإناث، الحياة والموت، والبيانات المرتبطة يمكن صيغتها كمصفوفة.

الأفراد	A1	A2	...	Ak	
S1					$B1 = \sum X_{ij}$
S2					B2
.					.
Sn					
					$N - \sum X_{ij}$

$$T_j = \sum X_{ij}$$

$$B_i = \sum X_{ij}$$

لاختبار الفرضية الصفرية

H0: جميع المعالجات متكافئة في فعاليتها

H1: هناك على الأقل معالجة تختلف في فعاليتها عن بقية المعالجات
نستخدم اختبار كوكران والذي يعطى بالعلاقة التالية:

$$Q = \frac{k(k-1) \sum T_j^2 - (k-1) N^2}{KN - \sum B_i^2} \sim \chi^2_{k-1}$$

القرار:

المحسوبة > الدرجة نقبل في ان نرفض الفرضية الصفرية.

المحسوبة < الدرجة نرفض الفرضية الصفرية.

مثال 5-12:

إذا كانت علامات 5 أشخاص في ثلاثة أسئلة موضوعية كما يلي:

الأفراد	Q1	Q2	Q3	Total
S1	1	1	0	2
S2	1	0	1	2
S3	0	0	1	1
S4	0	1	1	2
S5	1	0	1	2
	3	2	4	9

هل تدل هذه النتائج على اختلاف صعوبة الفقرات عند مستوى الدلالة الاحصائية

$$\alpha = 0.05$$

H0: الفرضية الصفرية: جميع الفقرات لها نفس درجة الصعوبة

H1: الفرضية البديلة: هناك على الأقل فقرة تختلف في صعوبتها عن بقية الفقرات

$$Q = \frac{k(k-1) \sum T_j^2 - (k-1) N^2}{KN - \sum B_i^2} \sim$$

$$Q = \frac{3(3-1)[32+22+42] - 2*9^2}{3*9 - [2^2+2^2+1^2+2^2+2^2]} = \frac{174 - 162}{27 - 17} = 1.2$$

$$\chi^2_{2,0.05} = 5.991 \quad \text{القيمة الحرجة}$$

القرار:

بما ان القيمة المحسوبة (1.2) > أقل من القيمة الحرجة ($\chi^2_{2,0.05} = 5.991$) نفشل في رفض الفرضية الصفرية، يعني الفقرات الثلاثة متساوية في درجة صعوبتها.

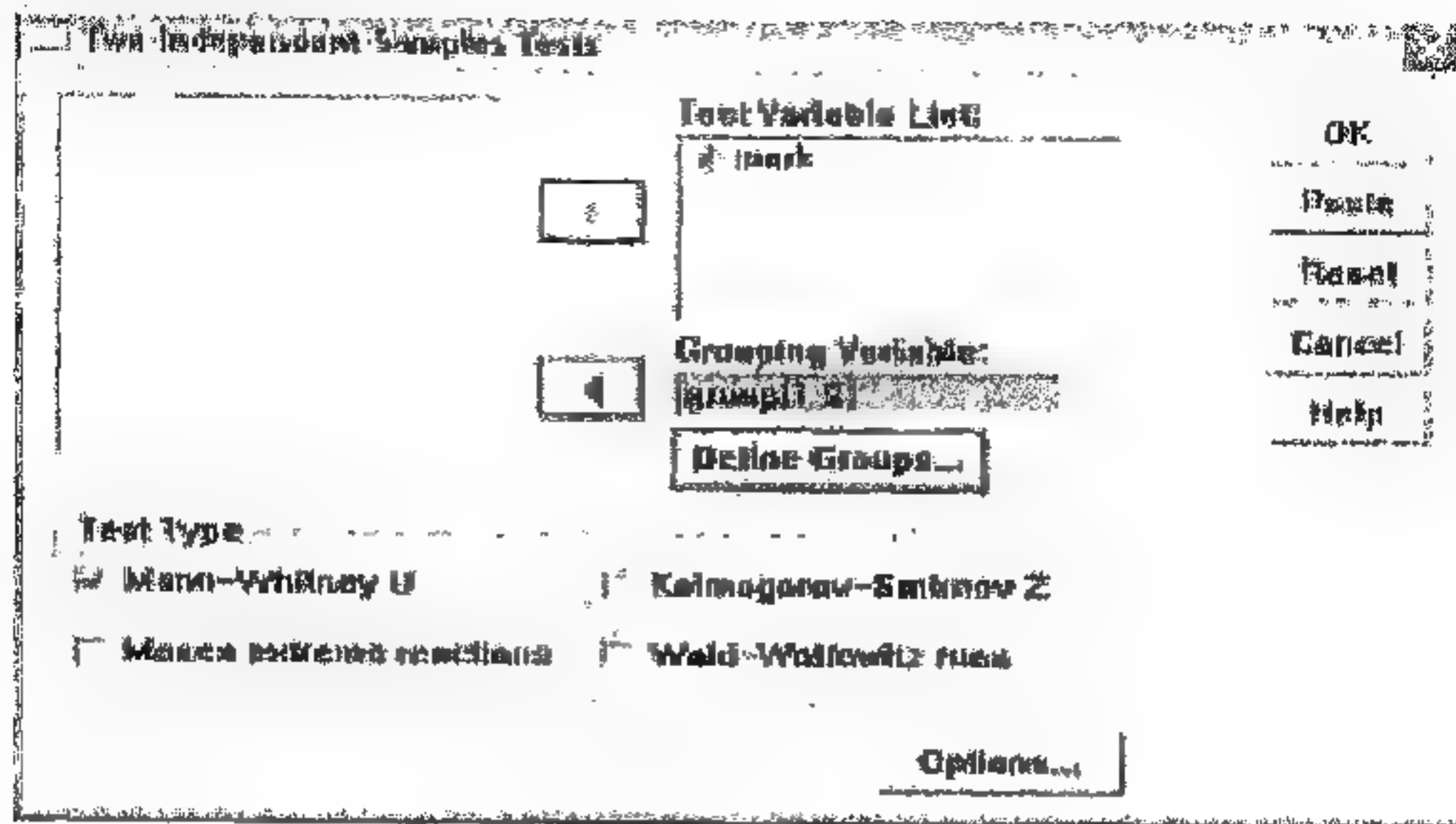
7-12 استخدام برمجية SPSS في حل المسائل.

مثال 2-12: نستخدم اختبار مان-وتني - عيتين مستقلتين.

* ادخال البيانات

group	mark
1	7.40
1	7.40
1	7.10
1	6.90
1	6.80
1	6.30
1	5.00
1	4.20
1	4.10
1	2.20
2	6.60
2	5.80
2	5.40
2	5.10
2	5.00
2	4.30
2	3.90
2	3.30
2	2.40
2	1.70

Analyze - Nonparametric Test - 2 Independent Samples...



ضع المتغير mark في نافذة Test Variable List، والمتغير group في نافذة Grouping Variable:
 من Test Type انقر مربع الاختبار Mann-Whitney U، ثم انقر زر Ok تظهر الشاشة أدناه:

Mann-Whitney Test

Ranks

	GROUP	N	Mean Rank	Sum of Ranks
MARK	1.00	17	17.44	296.50
	2.00	10	8.15	81.50
	Total	27		

Test Statistics^b

	MARK
Mann-Whitney U	26.500
Wilcoxon W	81.500
Z	-2.938
Asymp. Sig. (2-tailed)	.003
Exact Sig. [2*(1-tailed Sig.)]	.002 ^a

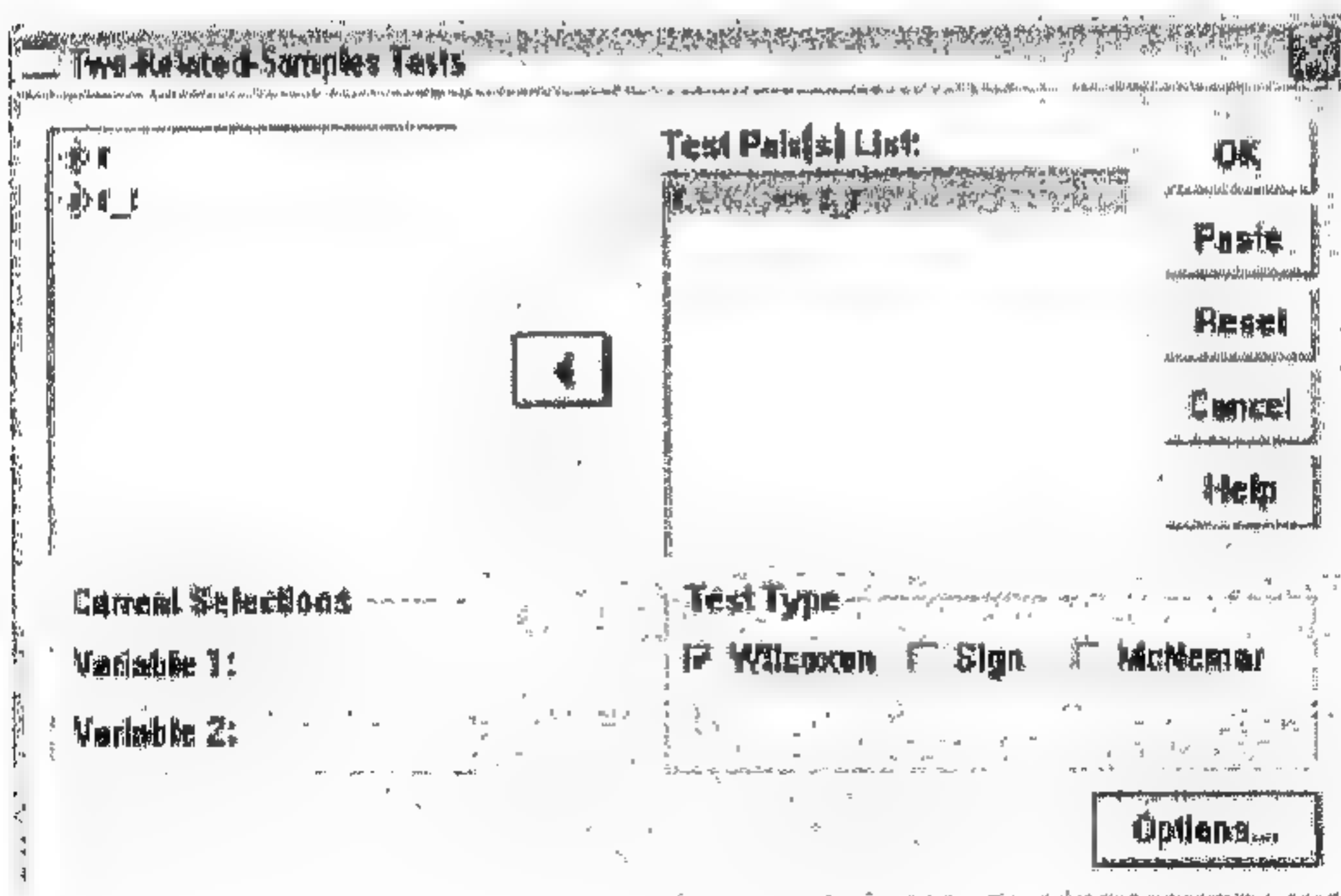
a. Not corrected for ties.

مثال 12-3: نستخدم اختبار ويلكسون - عيتين مترابطتين.

* ادخال البيانات

402	437
409	470
415	408
418	448
426	454
450	476
456	535
462	461
462	494
463	523

Analyze - Nonparametric Test - 2 Related Samples...



Test Pair(s) List: ضع المتغيران r , r_r في نافذة

من Test Type انقر مربع الاختبار Wilcoxon، ثم انقر زر Ok تظهر الشاشة أدناه:

NPar Tests

Wilcoxon Signed Ranks Test

Ranks

	N	Mean Rank	Sum of Ranks
R_H = R Negative Ranks	2 ^a	1.60	3.00
Positive Ranks	8 ^b	6.50	52.00
Ties	0 ^c		
Total	10		

a R_H < R

b R_H > R

c R = R_H

Test Statistics^a

	R_H = R
Z	-2.497 ^a
Asymp. Sig. (2-tailed)	.013

a. Based on negative ranks.

b. Wilcoxon Signed Ranks Test

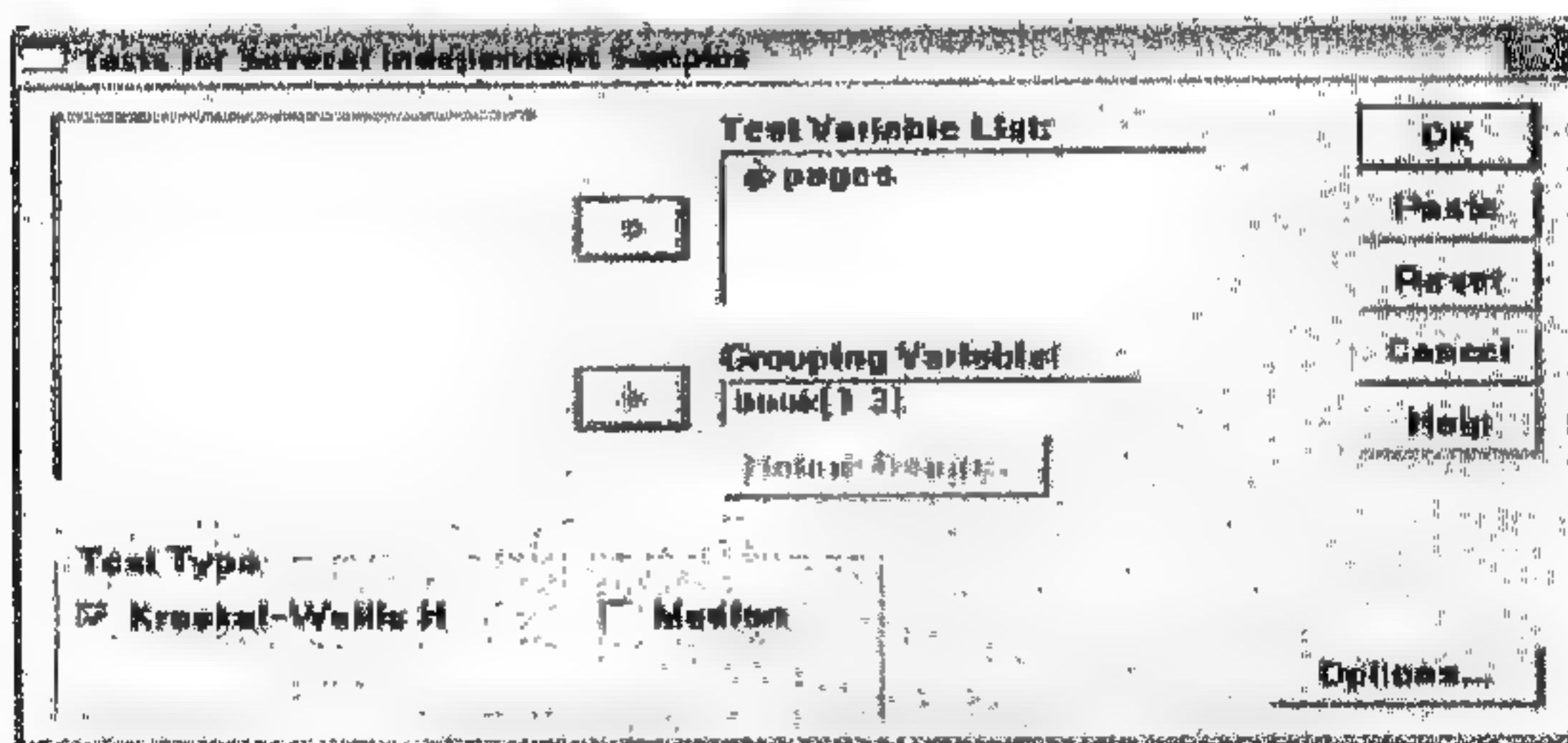
مثال 12-4: ستستخدم اختبار كروسكال والس - ثلاث عينات مستقلة أو أكثر

* ادخال البيانات

	book	pages
1	1.00	126.00
2	1.00	142.00
3	1.00	156.00
4	1.00	228.00
5	1.00	245.00
6	1.00	246.00
7	1.00	370.00
8	1.00	419.00
9	1.00	433.00
10	1.00	454.00
11	1.00	478.00
12	1.00	503.00
13	2.00	93.00
14	2.00	98.00
15	2.00	216.00
16	2.00	249.00
17	2.00	301.00
18	2.00	319.00
19	2.00	731.00
20	2.00	910.00

21	3.00	29.00
22	3.00	39.00
23	3.00	60.00
24	3.00	78.00
25	3.00	82.00
26	3.00	112.00
27	3.00	125.00
28	3.00	170.00
29	3.00	192.00
30	3.00	224.00
31	3.00	263.00
32	3.00	275.00
33	3.00	276.00
34	3.00	286.00
35	3.00	369.00
36	3.00	756.00

Analyze - Nonparametric Test - K Independent Samples...



صع المتغير pages في نافذة Test Variable List، والمتغير book في نافذة

Grouping Variable:

من Test Type انقر مربع الاختبار Kruskal-Wallis H، ثم انقر زر Ok تظهر

الشاشة أدناه:

→ NPar Tests

Kruskal-Wallis Test

Ranks			
PAGES	BOOK	N	Mean Rank
	stat	12	22.50
	math	8	21.13
	comp	18	14.19
	Total	38	

Test Statistics ^{a,b}	
	PAGES
Chi-Square	4.907
df	2
Asymp. Sig.	.086

a. Kruskal Wallis Test

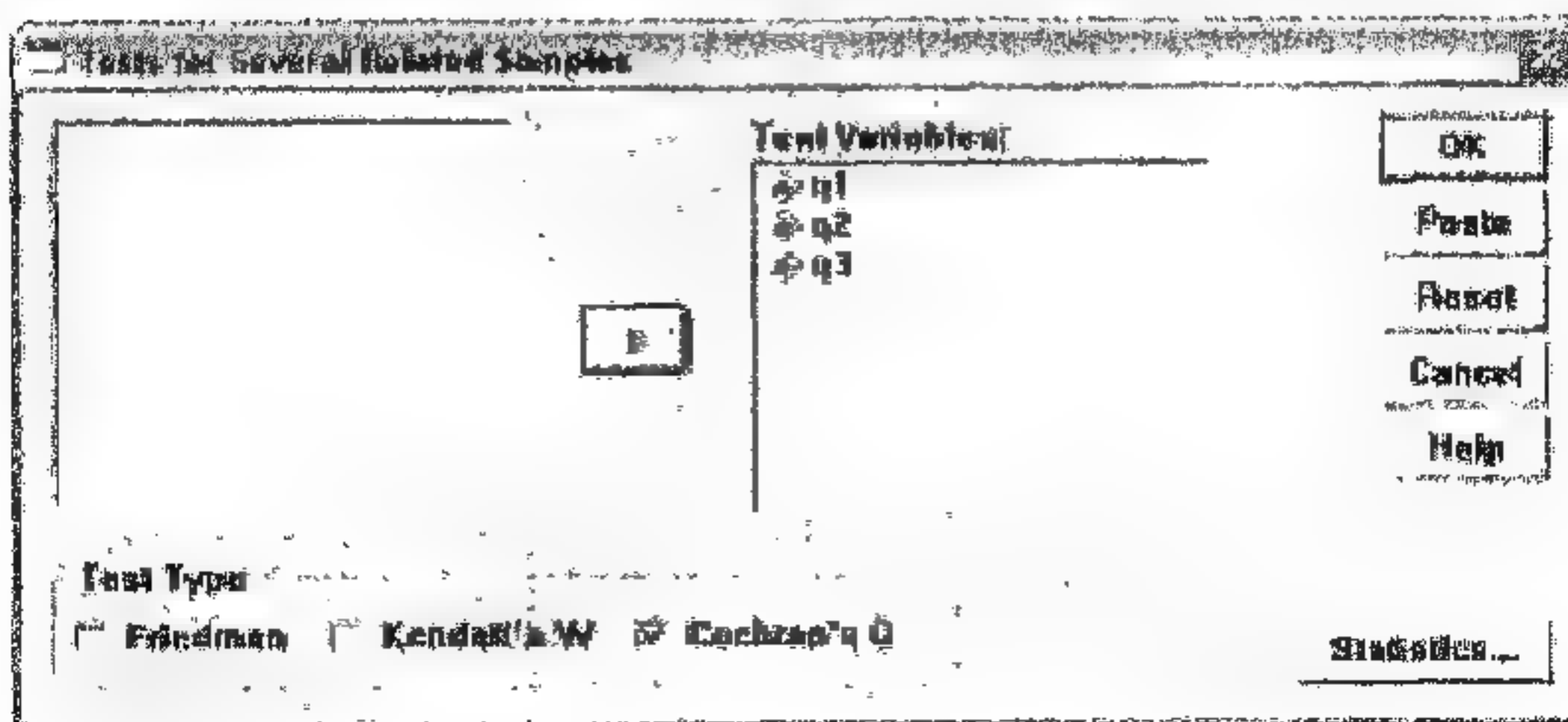
b. Grouping Variable: BOOK

القرار: بما أن القيمة المحسوبة $\chi^2 = 4.907 >$ أقل من القيمة الحرجة $\chi^2_{2,0.95} = 5.991$ نفسل في رفض الفرضية الصفرية، ونستنتج أن مجموعات الكتب الثلاث من نفس التوزيع من حيث عدد صفحات الكتاب، أي أن لها نفس الوسيطات ولها نفس التشتت ونفس الشكل.

مثال 12-5: نستخدم الطرق اللامعلمية (ثلاثة عينات مرتبطة أو أكثر)* اختبار كوكران * ادخال البيانات

q1	q2	q3
1	1	0
1	0	1
0	0	1
0	1	1
1	0	1

Analyze - Nonparametric Test - K Related Samples...



Test Variables: ضع المتغيران q1, q2, q3 في نافذة

من Test Type أنقر مربع الاختبار Cochran's Q، ثم انقر زر Ok تظهر الشاشة أدناه:

→ NPar Tests

Cochran Test

Frequencies

	Value	
	0	1
Q1	2	3
Q2	3	2
Q3	1	4

Test Statistics

N	5
Cochran's Q	1.200 ^a
df	2
Asymp. Sig.	.549

a. 1 is treated as a success.

8-12 تمارين Exercise.

س1: التالية علامات 15 طالب في مساق مهارات الحاسوب قبل وبعد تلقيهم لدوره في

الحاسوب، هل للدورة التدريبية أثر في تحسن هؤلاء الطلاب؟

* استخدم اختبار ويلكسون للأزواج المترابطة.

الرتب ذات الإشارة الأقل تكراراً	الرتب للفرق	الفرق = بعدى - قبلي	العلامة بعد الدورة	العلامة قبل الدورة	رقم الطالب
			70	65	1
			45	40	2
			55	60	3
			82	65	4
			60	80	5
			50	25	6
			63	40	7
			50	30	8
			83	85	9
			72	75	10
			55	45	11
			70	70	12
			88	74	13
			65	35	14
			80	65	15

Wilcoxon Signed Ranks Test

Ranks

	N	Mean Rank	Sum of Ranks
POST - PRE Negative Ranks	4 ^a	4.38	17.50
Positive Ranks	10 ^b	8.75	87.50
Ties	1 ^c		
Total	15		

a. POST < PRE

b. POST > PRE

c. PRE = POST

Test Statistics^b

	POST - PRE
Z	-2.200 ^a
Asymp. Sig. (2-tailed)	.028

a. Based on negative ranks.

b. Wilcoxon Signed Ranks Test

س2: اراد باحث مقارنة مجموعة تجريبية مع أخرى ضابطة في ادائهما على مقياس للاتجاهات، وكانت درجاتهم كما هو مبين أدناه:

* استخدم اختبار مان-وتني لعينتين مستقلتين.

المجموعة التجريبية Experimental	الرتبة Rank	المجموعة الضابطة Control	الرتبة Rank
52		52	
68		39	
42		47	
49		38	
36		27	
31		18	
29		20	
28		15	
50			

Mann-Whitney Test

Ranks

GROUP	N	Mean Rank	Sum of Ranks
MARK experimental	9	10.72	96.50
Control	8	7.06	56.50
Total	17		

Test Statistics^b

	MARK
Mann-Whitney U	20.500
Wilcoxon W	56.500
Z	-1.492
Asymp. Sig. (2-tailed)	.136
Exact Sig. [2*(1-tailed Sig.)]	.139 ^a

a. Not corrected for ties

b. Grouping Variable: GROUP

س3: نفترض أن شركة قامت باستخدام ثلاثة برامج لتدريب الموظفين، وبعد الانتهاء قدم بإجراء قياسات رتبية على أدؤهم في العمل، والجدول أدناه يبين علاماتهم على المقياس المستخدم، والمطلوب معرفة إن كان هناك اختلاف فيما بين المجموعات يعزى للأسلوب التدريبي.

* استخدم اختبار كرسكال والس.

Group_1	Group_2	Group_3
75	70	60
70	60	50
80	50	45
65	40	50
		49

→ NPar Tests

Kruskal-Wallis Test

Ranks

	GROUP	N	Mean Rank
MARK	1.00	4	11.13
	2.00	4	6.00
	3.00	5	4.50
	Total	13	

Test Statistics^{a,b}

	MARK
Chi-Square	6.926
df	2
Asymp. Sig.	.031

a. Kruskal Wallis Test

b. Grouping Variable: GROUP

س4: ضع رمز الإجابة الصحيحة (نعم، لا) في المربع المخصص لذلك:

الرقم	1	2	3	4	5	6	7	8	9	10
الإجابة										

- 1- تتطلب الاختبارات اللامعلمية الإيفاء بافتراضات تجانس التباين.
- 2- تتطلب الاختبارات اللامعلمية الإيفاء بافتراضات التوزيع الطبيعي.
- 3- الاختبارات المعلمية أكثر قوة من الاختبارات اللامعلمية.
- 4- تميل الاختبارات اللامعلمية لرفض الفرضية الصفرية.
- 5- يعتبر اختبار ولكوكسون بديلاً لاختبار (T) للبيانات المستقلة.
- 6- يعتبر اختبار مان-وتني بديلاً لاختبار (T) للبيانات المترابطة.
- 7- يعتبر اختبار كروسكال والس بديلاً لتحليل التباين الأحادي.

الفصل الثالث عشر

تطبيقات البرنامج الإحصائي SPSS

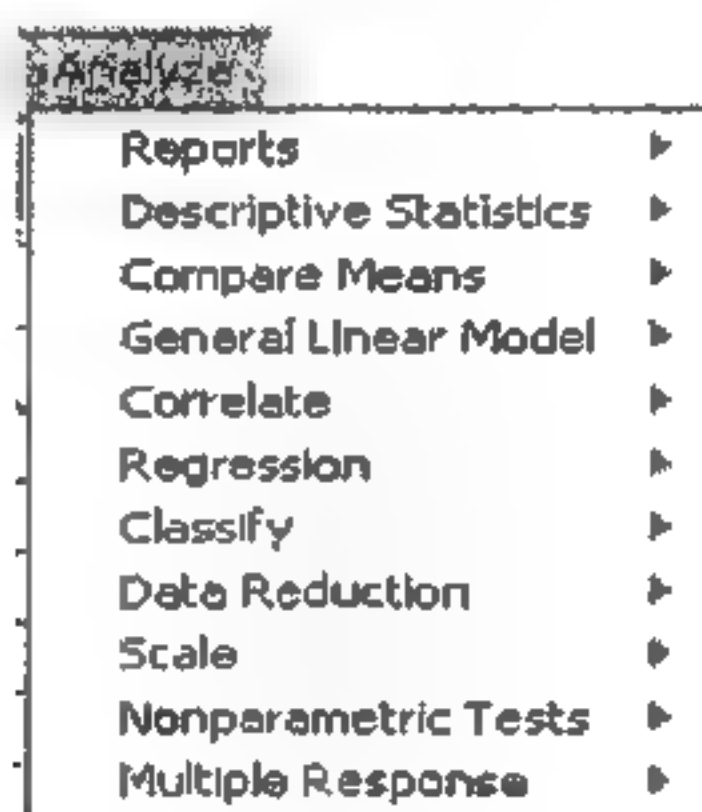
- 1-13 التعرف على بيئة النظام الإحصائي SPSS
- 2-13 تشغيل نظام SPSS.
- 3-13 شاشات نظام SPSS.
- 4-13 ملفات نظام SPSS.
- 5-13 القوائم الرئيسية لبرنامج SPSS.
- 6-13 استخدام برمجية SPSS في حل المسائل.

الفصل الثالث عشر

تطبيقات البرنامج الإحصائي SPSS

1-13 التعرف على بيئة النظام الإحصائي SPSS

* أن كلمة SPSS تتكون من Statistical Package for Social Sciences وهي تعني الحزمة الإحصائية للعلوم الاجتماعية، وتستخدم لأجراء عمليات إحصائية كثيرة وبشكل سهل. وتوجد أغلب التحليلات في قائمة التحليل Analyze المبينة أدناه:



يلزمنا أن نقوم بإجراء بعض التحليلات الإحصائية للبيانات مثل: الإحصاء الوصفي Descriptive Statistics، مثل إيجاد مقاييس التزعة المركزية: الوسط الحسابي، الوسيط، المنوال، ... وإيجاد مقاييس التشتت: المدى، التباين، الانحراف المعياري، المدى المتوسط، ... مقارنة الأوساط Compare Means.

إيجاد معاملات الارتباط Correlate: الارتباط البسيط، الارتباط المتعدد، الارتباط الجزئي، الارتباط شبه الجزئي،

إيجاد معادلة الانحدار Regression

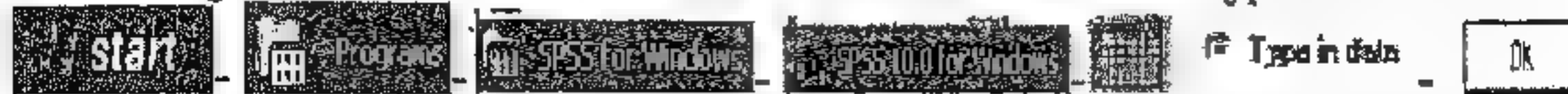
الإحصاءات اللامعلمية Nonparametric Tests

إن القيام بالحسابات اليدوية لهذه الإحصاءات ليس سهلاً خاصة إذا كان حجم البيانات كبيراً، ولكن باستخدام برنامج النظام الإحصائي SPSS تصبح جميع هذه التحليلات سهلة وممتعة إذا تم التعامل مع البرنامج بشيء من الحرفية.

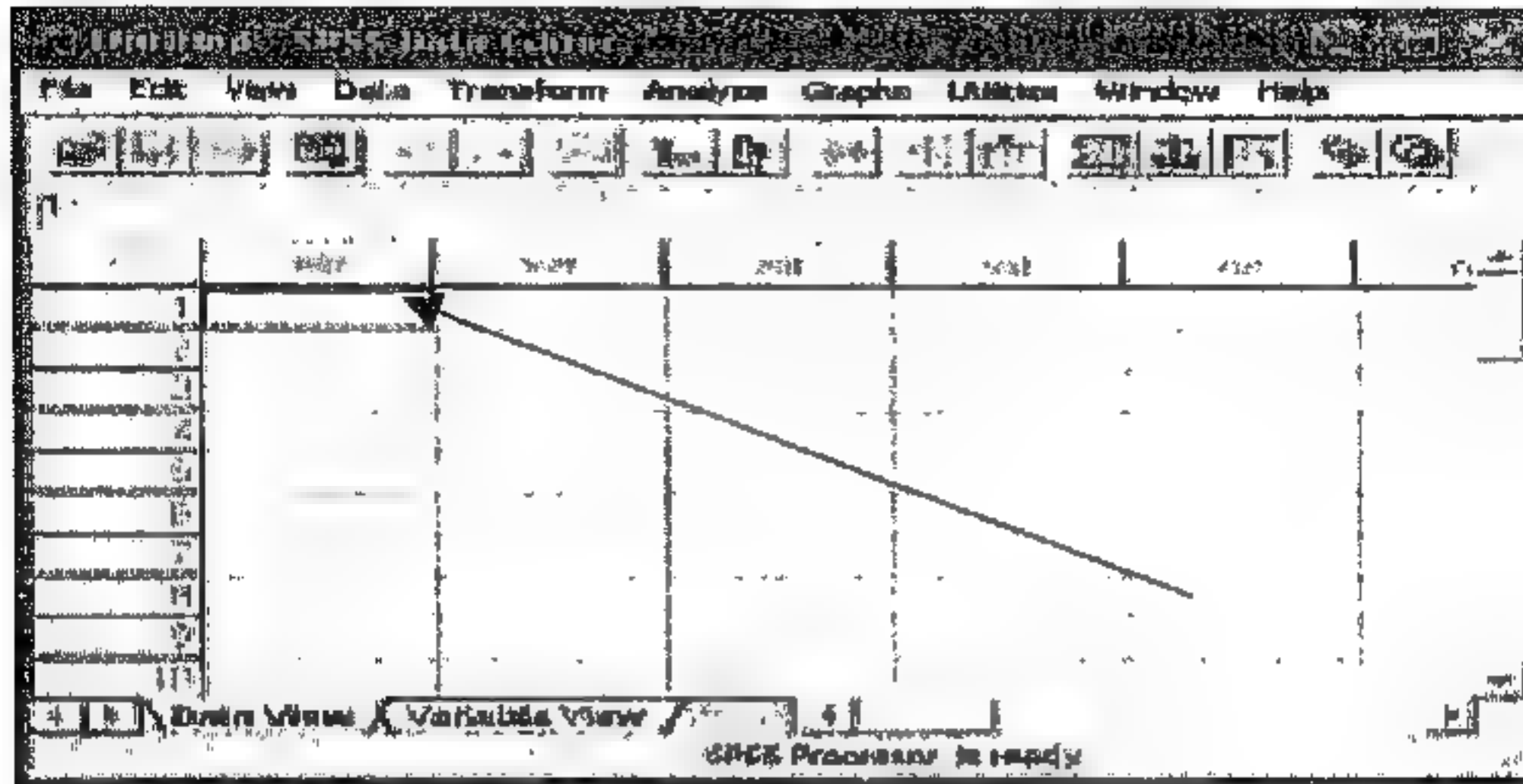
2-13 تشغيل نظام SPSS.

* تشغيل نظام SPSS

Start - Programs - SPSS for Windows - SPSS10.0 for Windows - Type in data Ok



تظهر لديك الشاشة المبينة أدناه:



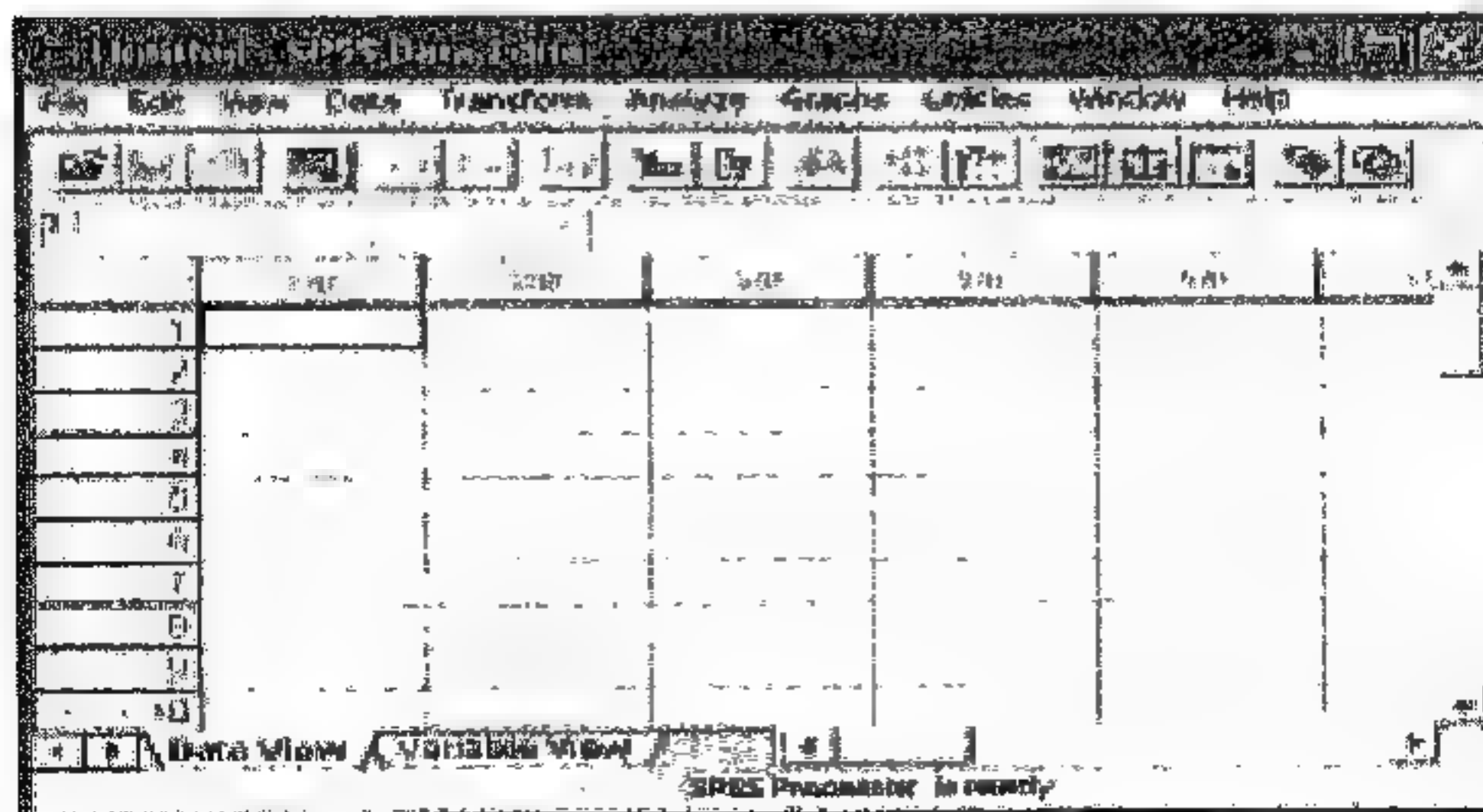
3-13 شاشات نظام SPSS:

يحتوي نظام SPSS على ثلاث شاشات رئيسية هي:

1. شاشة محرر البيانات Data Editor Window: وتحتوي على البيانات الإحصائية

المراد تحليلها.

وتتكون من شاشتين هما Data View وتستخدم لإدخال البيانات

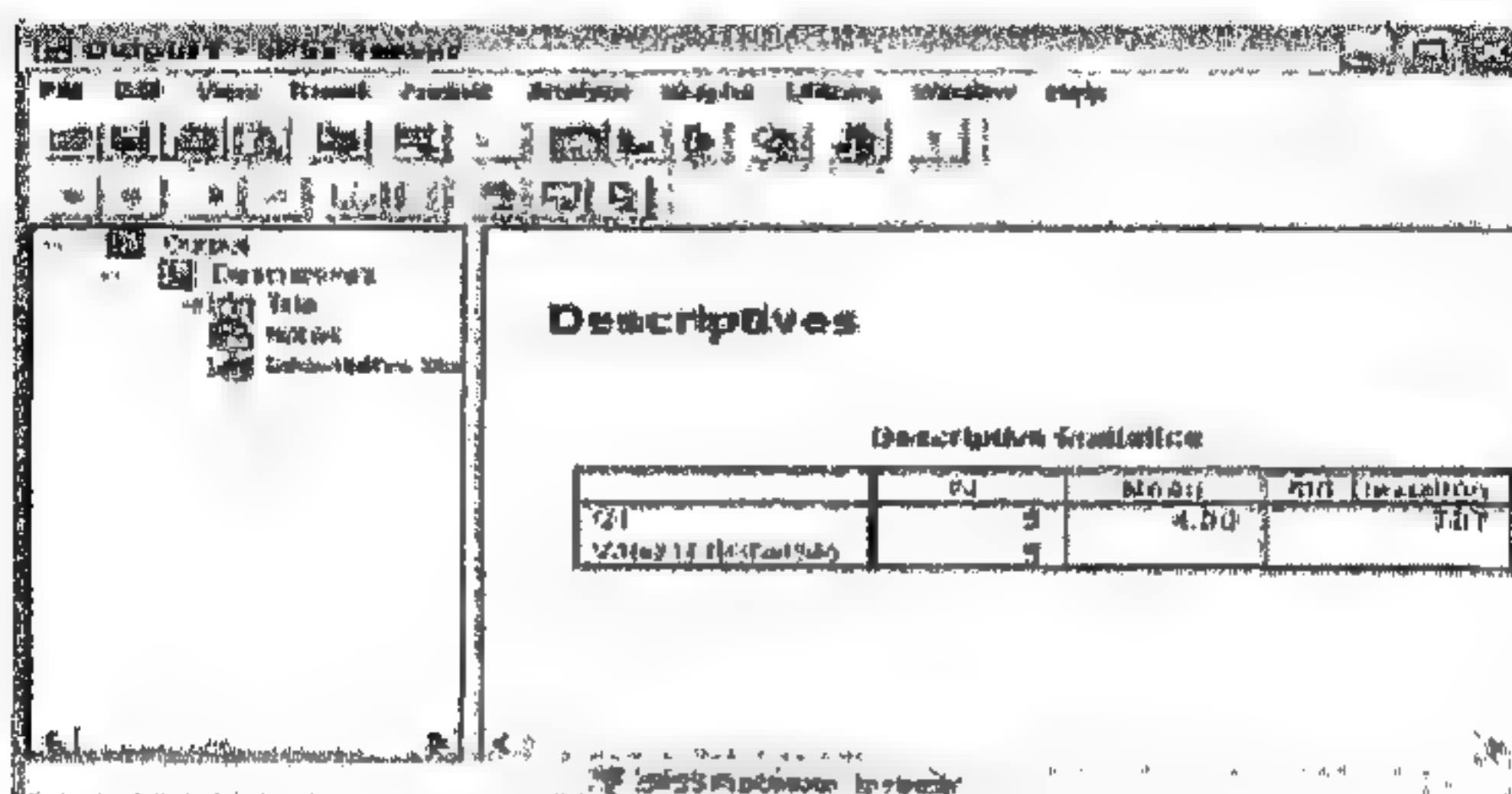


وشاشة Variable View وهي شاشة تعريف المتغيرات كما هو مبين أدناه:

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1										
2										
3										

2. شاشة المخرجات Output Navigator: تظهر نتائج التحليلات الإحصائية

والرسوم البيانية.

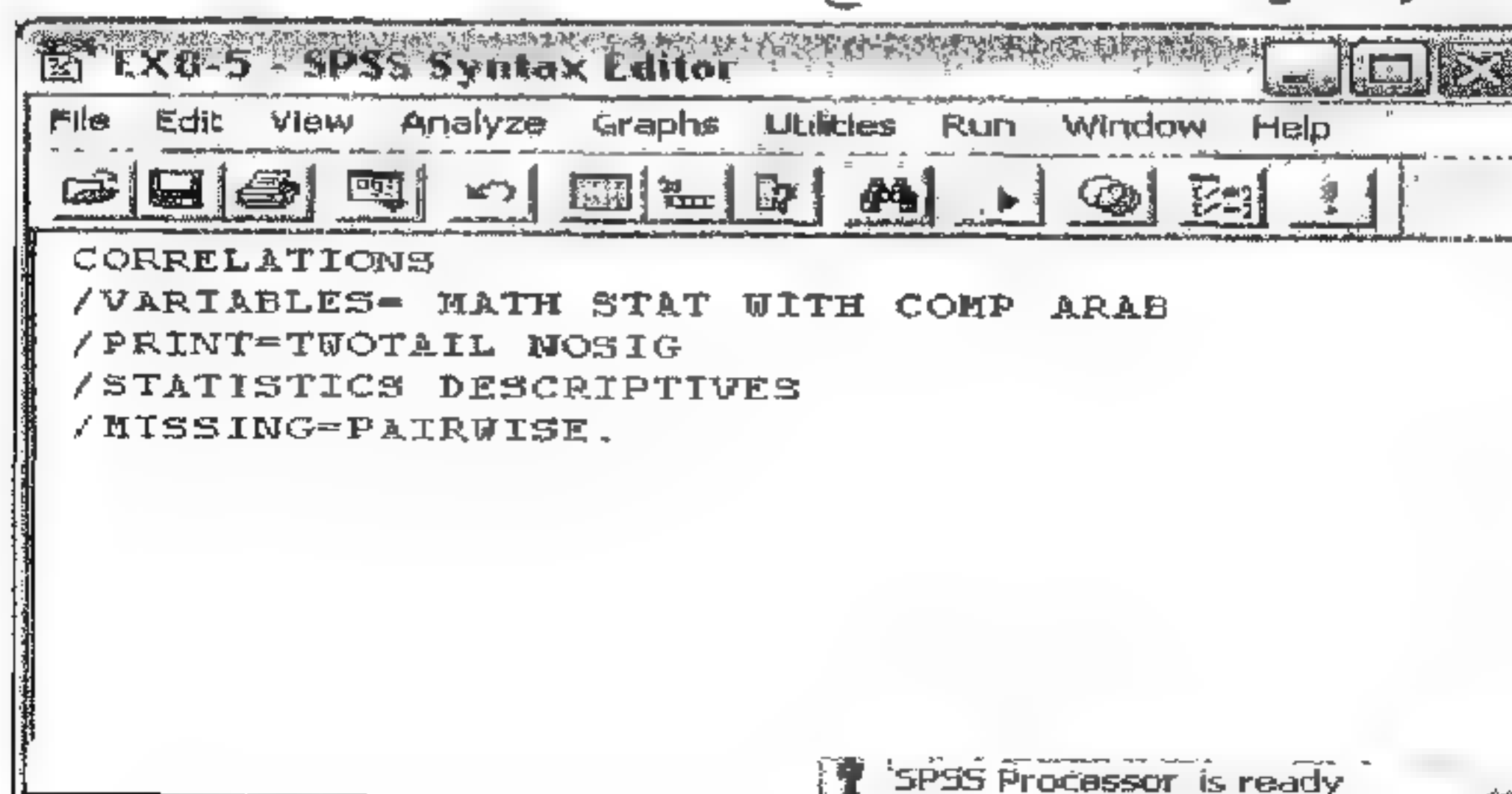


3. شاشة التعليمات Syntax Window: تستخدم لكتابة التعليمات للعمليات

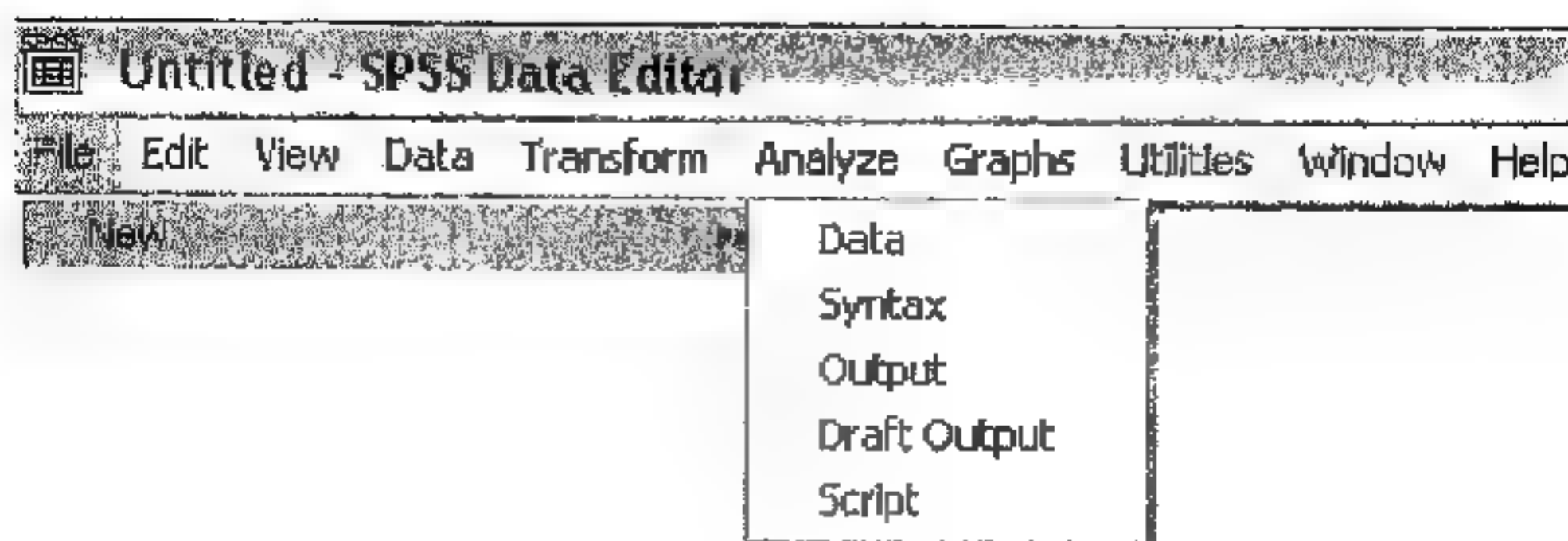
المختلفة.

من قائمة File اختر New ثم Syntax تظهر لك شاشة التعليمات أدناه، ثم اكتب

التعليمات في داخل الشاشة كما هو موضح أدناه:



4-13 ملفات نظام SPSS:



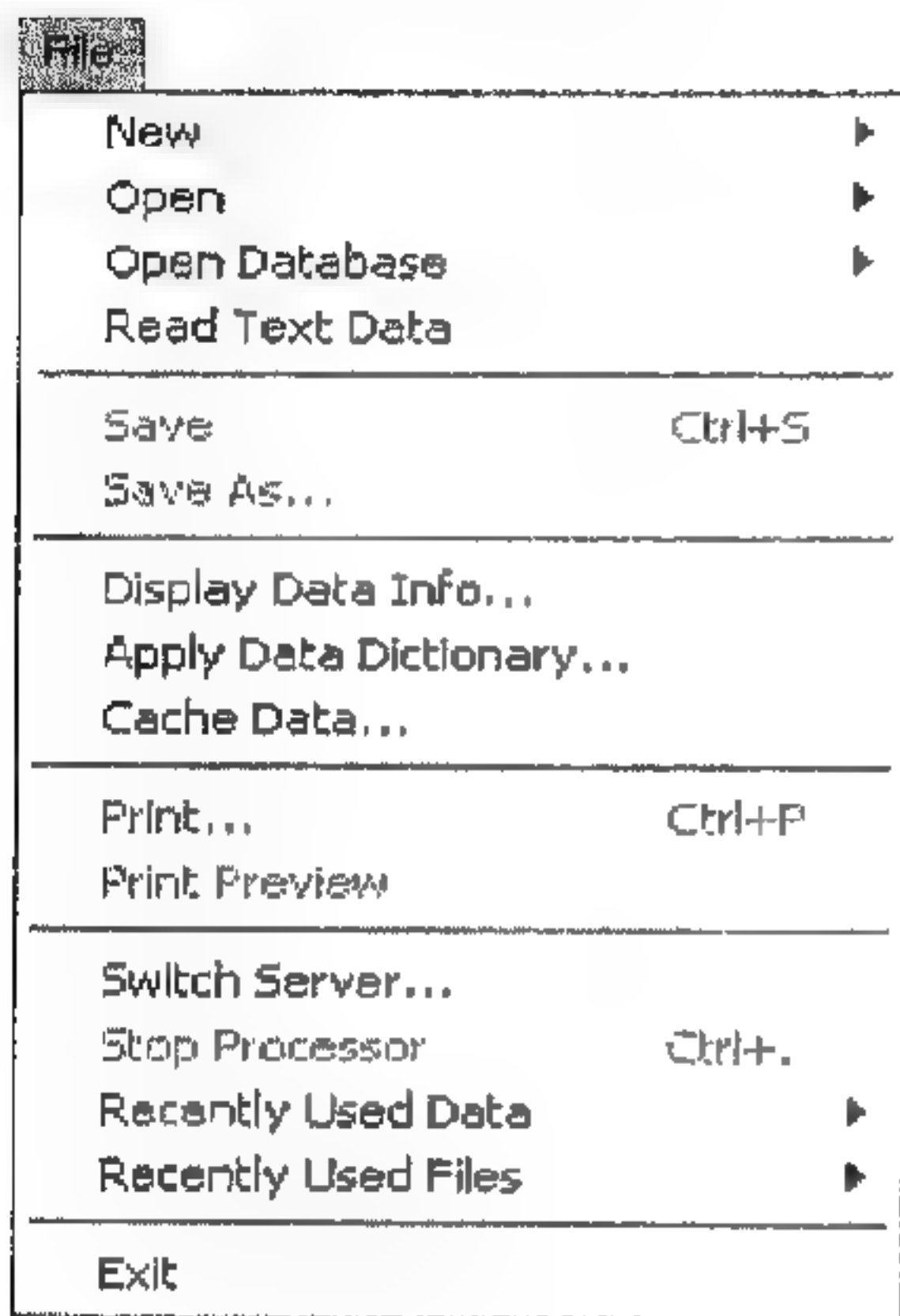
هناك عدة أنواع من الملفات منها ما يلي:

1. ملف البيانات **Data**: وهي الملفات التي تحتوي على البيانات الخام المراد تحليلها، ويكون نوع الملف (.SAV)، فإذا كان اسم ملف البيانات JAMAL فإن نوعه (.SAV)، فيكون الاسم الكامل للملف هو (JAMAL.SAV)
2. ملف المخرجات **Output**: وهو الملف الذي يحتوي على نتائج الإجراءات الإحصائية التي تظهر في شاشة المخرجات، ويكون نوع الملف (.SPO)، فإذا كان اسم ملف المخرجات JAMAL فإن نوعه (.SPO)، فيكون الاسم الكامل للملف هو (JAMAL.SPO)
3. ملف التعليمات **Syntax**: وهو الملف الذي يحتوي على التعليمات المراد إجراؤها، ويكون نوع الملف (.SPS)، فإذا كان اسم ملف التعليمات JAMAL فإن نوعه (.SPS)، فيكون الاسم الكامل للملف هو (JAMAL.SPS)
4. ملف **Draft Output**: وهو الملف الذي يحتوي على مسودة ملف المخرجات المراد إجراؤها، ويكون نوع الملف (.rtf)، مثال ذلك هو (JAMAL.rtf)
5. ملف **Script**: وهو الملف الذي يحتوي على الإجراءات الجاهزة وتستخدم لغة Sax Basic والتي تكتب بها الإجراءات الجاهزة مثل Clean Viewer ويكون نوع الملف (.Sbs)، فإذا كان اسم ملف الإجراءات Remove Labels فإن نوعه (.sbs)، فيكون الاسم الكامل للملف هو (Remove Labels.sbs).

5-13 القوائم الرئيسية لبرنامج SPSS:



قائمة ملف File



New: إنشاء ملفات جديدة.

Open: فتح ملفات مخزنة سابقاً.

Read Text Data: قراءة ملف بيانات.

Save: تخزين ملفات.

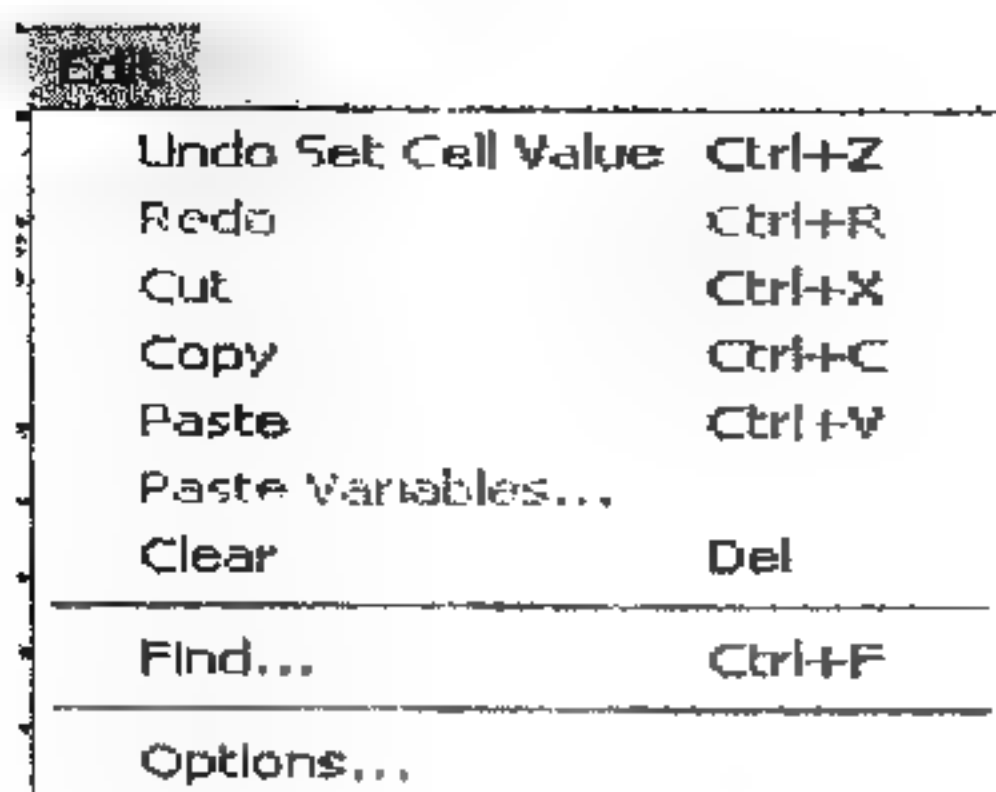
Display Data Information: إظهار معلومات عن الملفات.

Print: طباعة الملفات.

Print Preview: معاينة الملفات قبل الطباعة.

Exit: الخروج من البرنامج.

قائمة تحرير Edit



Undo Set Cell Value: التراجع عن التحرير.

Redo Set Cell Value: التراجع عن التراجع عن التحرير.

Cut: قص البيانات.

Copy: نسخ البيانات.

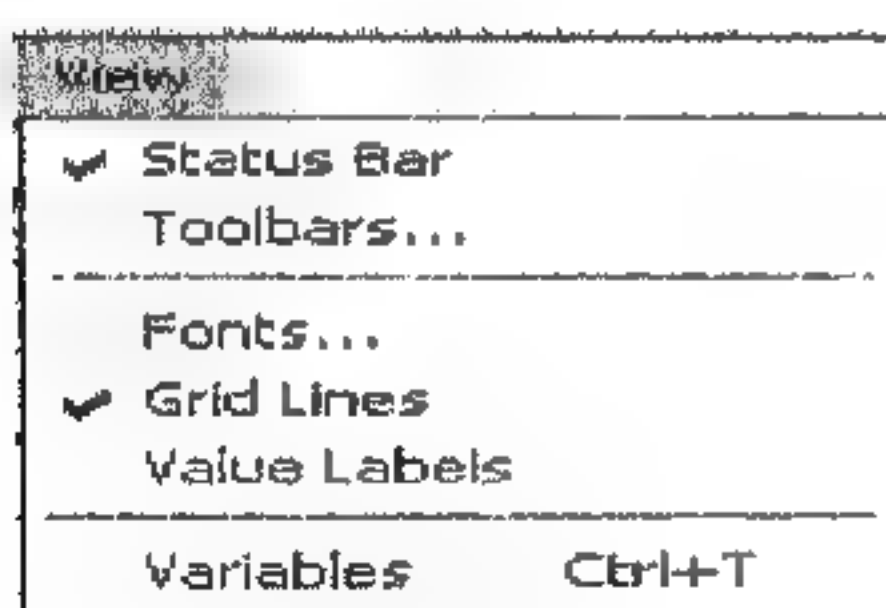
Paste: لصق البيانات.

Clear: حذف (عمود، صف) أو أكثر بما يحويه من بيانات بعد تحديد المراد حذفه.

Find: البحث عن حالات.

Options...: خيارات.

قائمة عرض View



Status Bar: إظهار أو إخفاء شريط الحالة.

Toolbars: إظهار شريط الأدوات والأيقونات المختصرة المستخدمة بدل من القوائم.

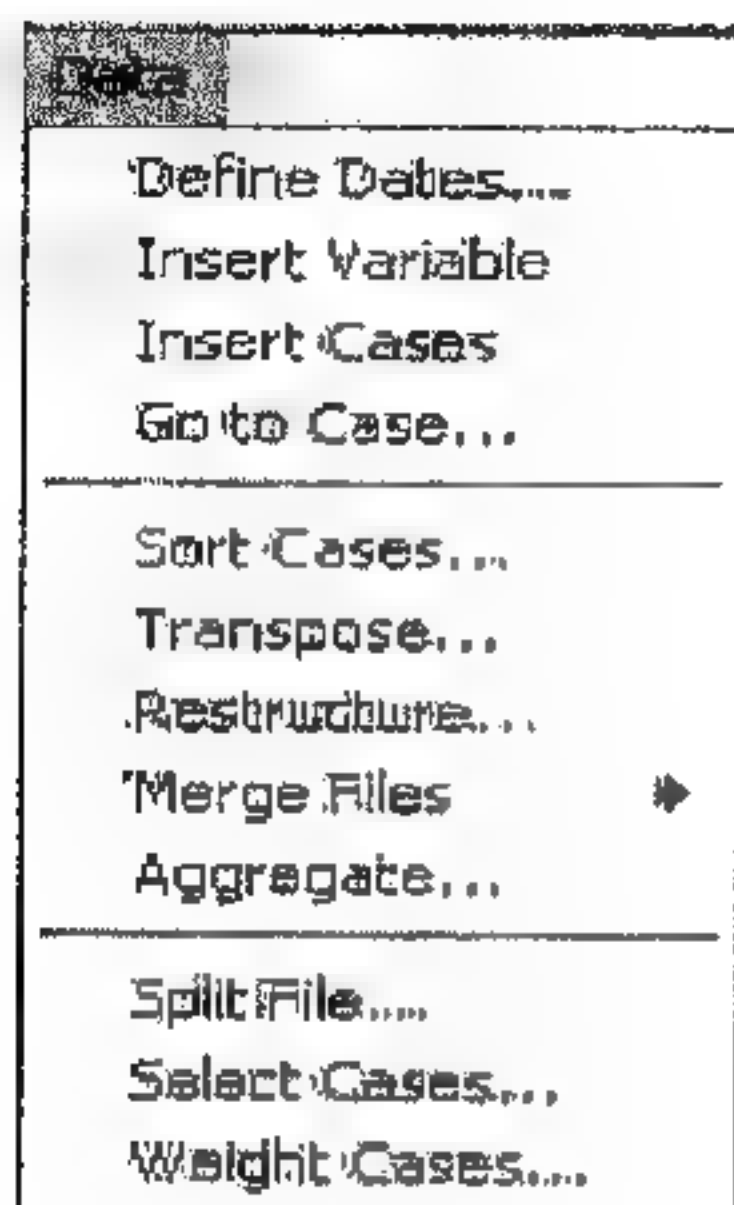
Fonts...: تغيير نوع الخط المستخدم.

Grid Lines: إظهار أو إخفاء خطوط الشبكة.

Value Labels: إظهار أو إخفاء عناوين القيم.

Variables: إظهار شاشة تعريف المتغيرات.

قائمة البيانات Data



Define Dates...: تعريف المتغيرات وتغيير اسمائها.

Insert Variables: اضافة متغيرات جديدة.

Insert Cases: اضافة حالات جديدة.

Go to Case...: الذهاب الى حالة معينة.

Sort Cases...: ترتيب الملف حسب قيم متغير ما.

Transpose...: تحويل البيانات.

Merge Files: دمج الملفات وهي دمج اكثر من ملف وجعلها ملف واحد.

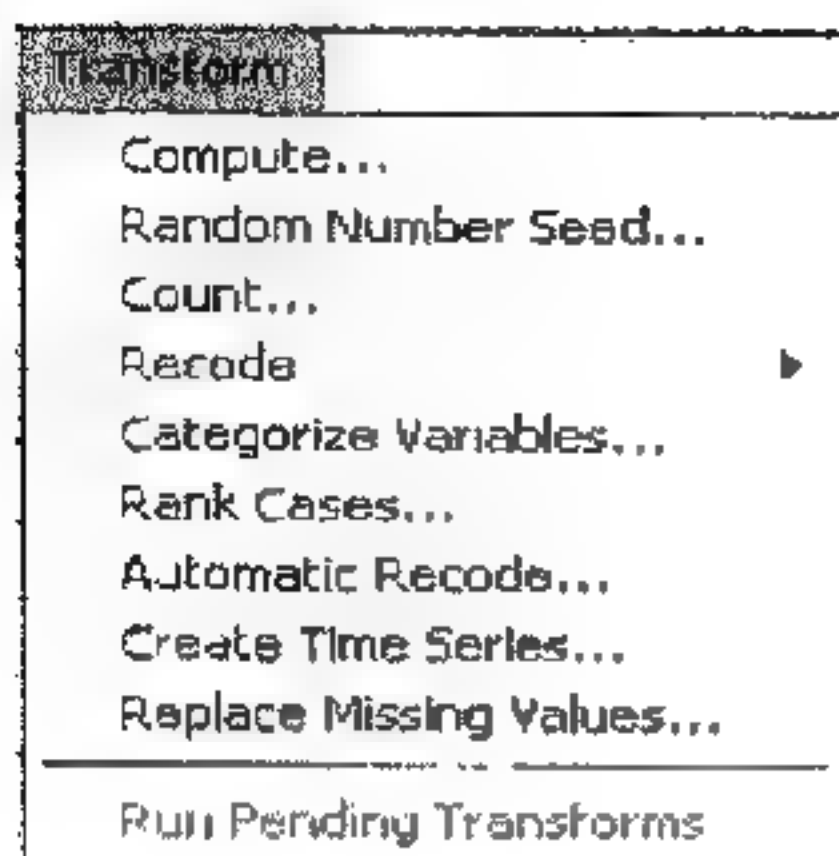
Aggregate...: تجميع وتلخيص الحالات.

Split File...: تقسيم الملف الى عدة أجزاء.

Select Cases...: تستخدم لاختيار مجموعة من الحالات ينطبق عليها شرط معين.

قائمة التحويلات Transform:

إنشاء متغير جديد من خلال المتغيرات الموجودة.



Compute: القيام بالعمليات الحسابية المختلفة.

Count : حساب عدد القيم المتشابهة لقائمة من المتغيرات لكل فرد من أفراد العينة.

Recode: إعادة ترميز البيانات.

Rank Cases...: إنشاء متغيرات جديدة تحتوي على رتب المتغيرات الموجودة المختلفة

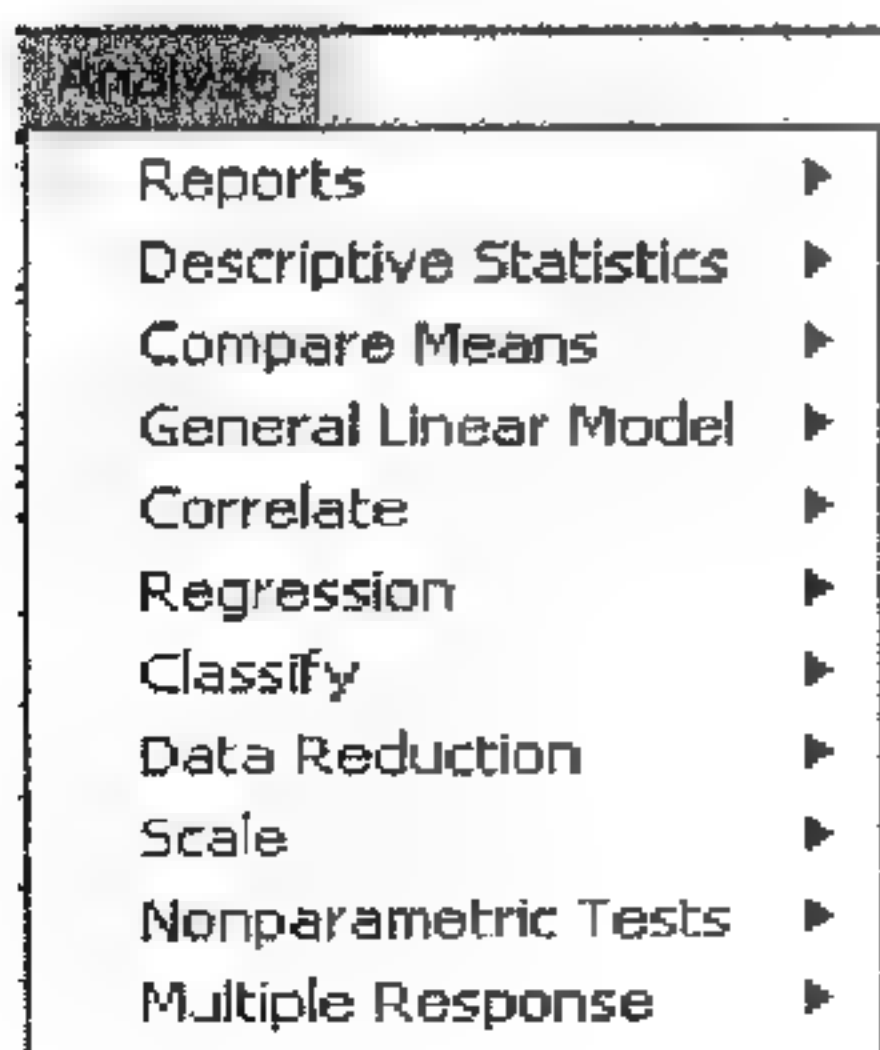
لقيم الرقمية.

Automatic Recode...: (الترميز الآلي)، إعادة ترميز السلاسل الحرفية إلى قيم.

Create Time Series...: إنشاء متغير جديد يحتوي متسلسلة زمنية.

Replace Missing Values...: تعويض القيم المفقودة بطرائق إحصائية.

قائمة التحليلات Analyze



Reports: التقارير.

Descriptive Statistics: الإحصاء الوصفي.

Compare Means: مقارنة الأوساط، تحليل التباين الأحادي.

General Linear Model: تحليل التباين الثنائي Tow-Way ANOVA

Correlate: حساب معاملات الارتباط.

Regression: حساب معادلة الانحدار.

Classify: التصنيف.

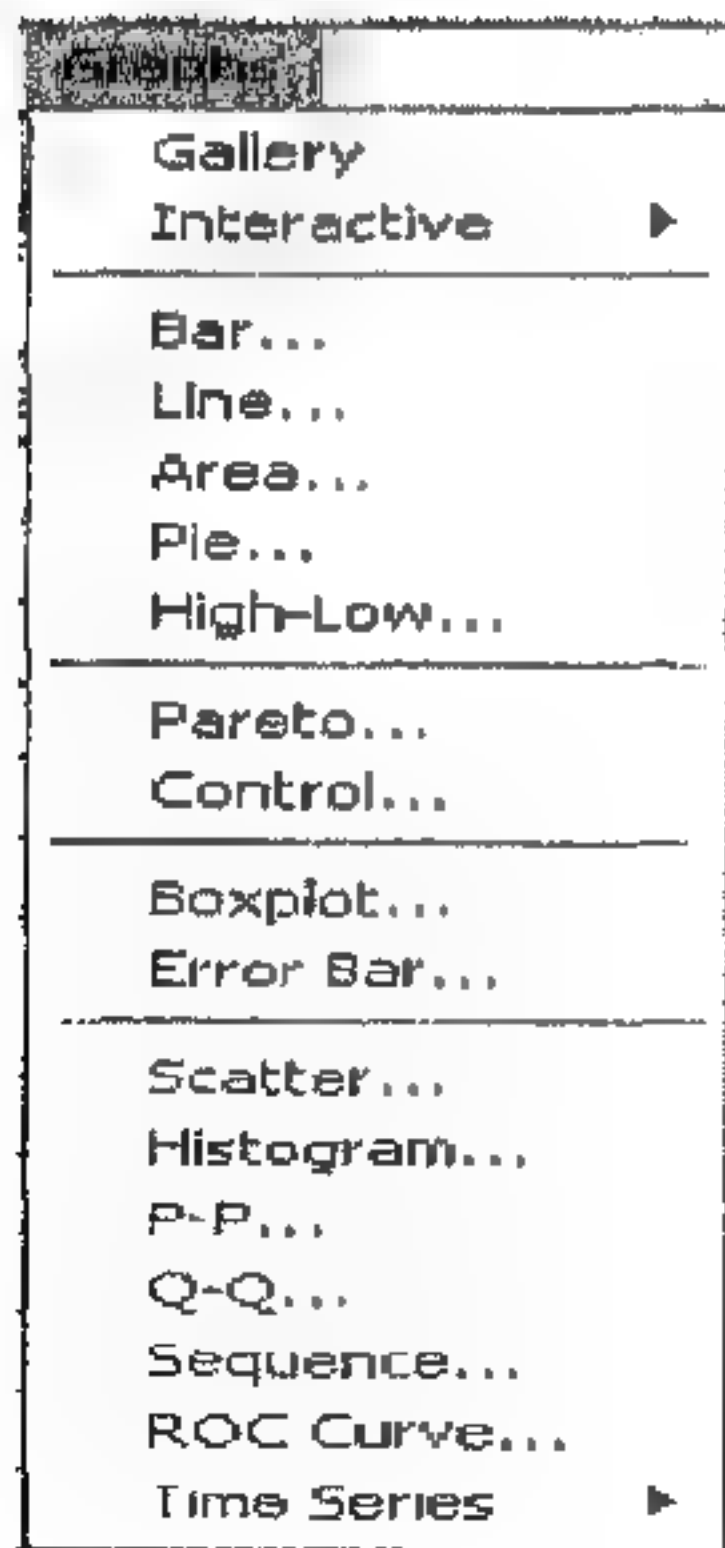
Data Reduction: التحليل العاملي.

Scale: تحليل الثبات (معاملات الثبات).

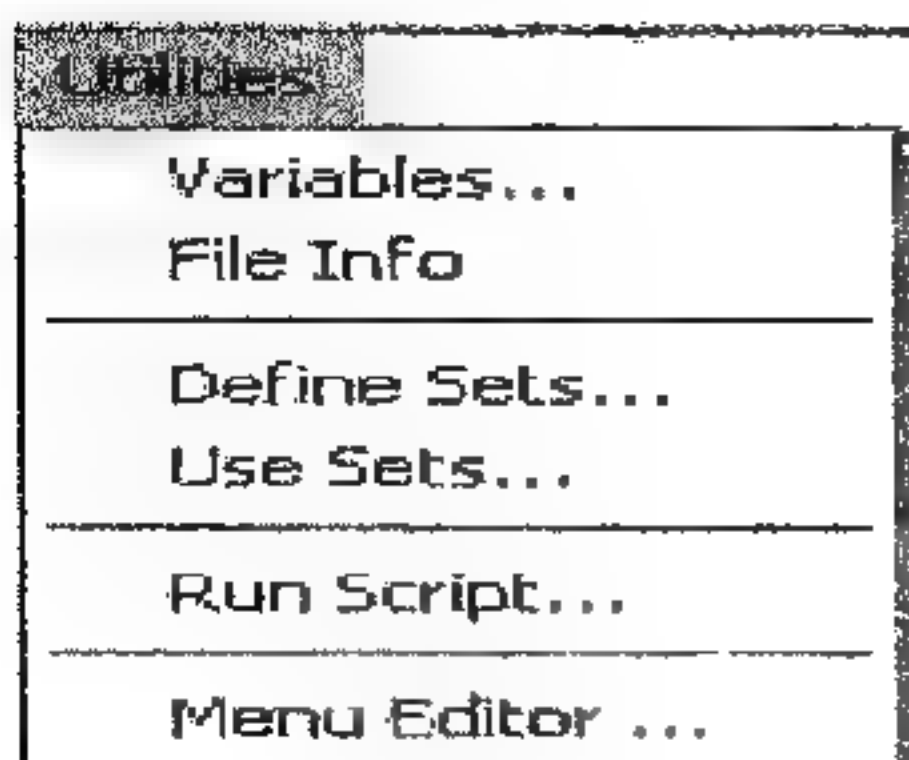
Nonparametric Test: الإحصاءات اللامعلمية.

Multiple Response: تعريف المجموعات.

قائمة الرسوم البيانية Graphs



قائمة الفوائد Utilities



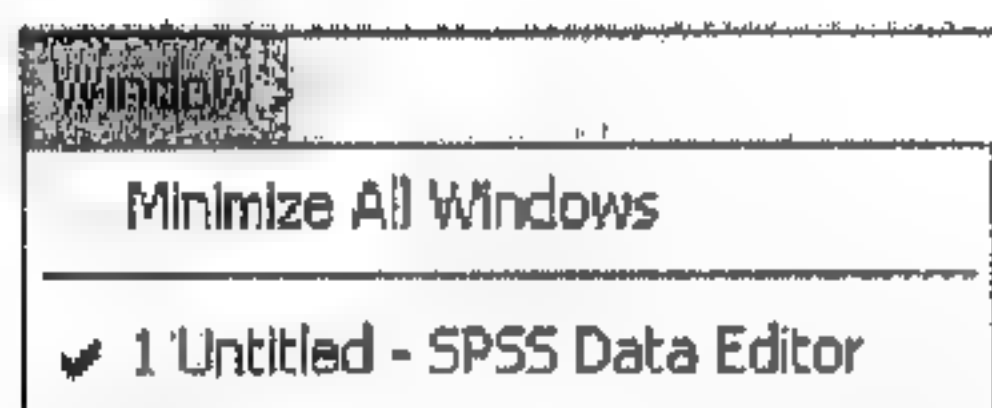
Variables... : إعطاء معلومات عن المتغيرات.

File Info : إيجاد معلومات مفصلة عن الملف المستخدم والمتغيرات التي به.

Define Sets... : تعريف المجموعات للمتغيرات المختلفة.

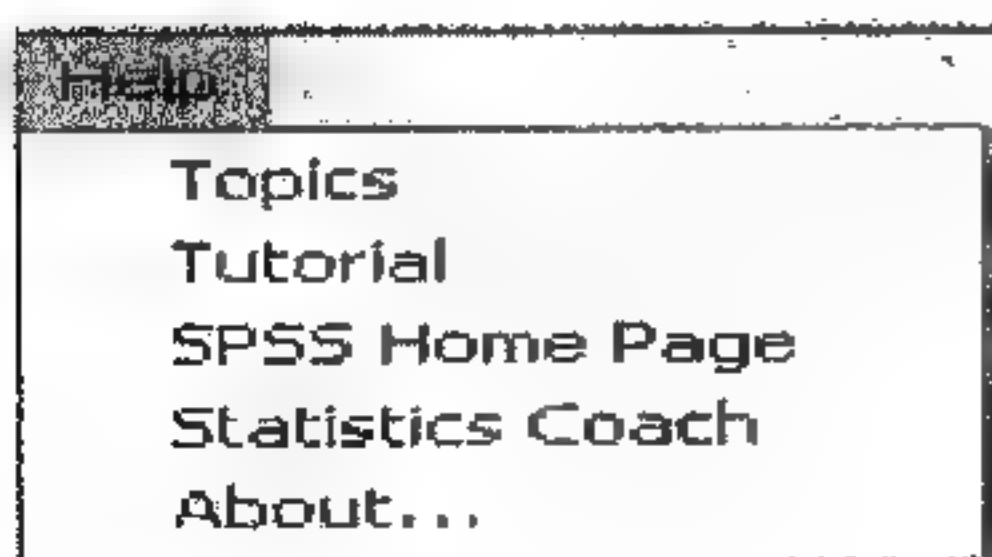
Use Sets... : استخدام المجموعات للمتغيرات المختلفة.

قائمة نافذة Window



Minimize All Windows : التحكم بحجم النوافذ.

قائمة المساعدة Help



Topics : إعطاء مساعدة عن أي محتوى من محتويات البرنامج.

6-13 استخدام برمجية SPSS في حل المسائل.

س2) في اختبار لمادة الحاسوب الذي يتكون من الأسئلة أدناه، حلل الاختبار؟

رقم الطالب: التخصص: 1-علمي 2-أدبي 3-معلوماتية 4-مهني

الجنس: 1-ذكر 2-أنثى .

المادة: الحاسوب. المستوى الثالث - تراسل البيانات والشبكات. (15 سؤال).

ضع دائرة حول رمز الإجابة الصحيحة في كل مما يلي:

1- أي من التالية ليس من مكونات شبكة الحاسوب:

- أ- أجهزة حاسوب مزودة ببطاقات خاصة بالاتصال. ج- منظم التيار الكهربائي.
- ب- أسلاك توصيل بين الأجهزة. د- بطاقات الشبكات.

2- تغطي الشبكة المحلية LAN مساحة قطرها حوالي:

- أ- 10م ب- 1000م ج- 100م د- 10000م

3- الجهاز الذي يعمل على زيادة حجم وكفاءة الشبكة هو:

- أ- المقسم Switch. ب- الخط Line. ج- الموزع Hub. د- الجسر Bridge.

4- شبكة مكونة من جهازي حاسوب فقط مرتبطين بخط اتصال تسمى:

- أ- الشبكة المحلية LAN. ج- شبكة الخادم والمستفيد Client/Server.
- ب- الشبكة الناعورية Peer-to-Peer. د- الشبكة النجمية Star Network.

5- إن المصطلح Hyper Text يعني:

- أ- نص مترابط. ب- عنوان الموقع. ج- مزود خدمة الانترنت. د- متصفح الانترنت.

6- إن المصطلح URL يعني:

- أ- نص مترابط. ب- عنوان الموقع. ج- مزود خدمة الانترنت. د- متصفح الانترنت.

7- إن المصطلح Search Engines يعني:

- أ- محركات البحث.
- ب- عنوان الموقع.
- ج- الشبكة العنكبوتية العالمية.
- د- متصفح الانترنت

8- أي من التالية ليست من محركات البحث Search Engines:

- أ- Yahoo
- ب- Google
- ج- Altavista
- د- Windows

9- يرمز للمواقع الحكومية بـ:

- أ- Edu
- ب- Org
- ج- Gov
- د- Net

10- تغطي الشبكة واسعة التغطية WAN مساحة تقدر بـ:

- أ- مدينة.
- ب- دولة.
- ج- قارة.
- د- الكرة الأرضية.

11- تقسم الشبكات المحلية LAN إلى:

- أ- الشبكة التناظرية Peer-to-Peer، الشبكة التماثلية Symetric Network
- ب- شبكة الخادم والمستفيد Client/Server، الشبكة التناظرية Peer-to-Peer
- ج- الشبكة التناظرية Peer-to-Peer، شبكة القيمة المضافة VAN.
- د- شبكة الخادم والمستفيد Client/Server، شبكة القيمة المضافة VAN.

12- تمتاز الشبكة التناظرية Peer-to-Peer بأنها:

- أ- مناسبة عندما تكون شبكة الاتصال كبيرة.
- ب- يوجد جهاز أساسي يؤدي الخدمات للشبكة وبقية الأجهزة محطات عمل.
- ج- تكون جميع الأجهزة في هذه الشبكة متكافئة.
- د- لكل جهاز حق الوصول إلى الشبكة حسب أهميته.

13- يتكون عنوان البريد الإلكتروني من:

- أ- اسم خاص بالمستخدم ، إشارة # ، اسم الموقع الذي يقدم خدمة البريد الإلكتروني.
- ب- اسم خاص بالمستخدم ، إشارة \$ ، اسم الموقع الذي يقدم خدمة البريد الإلكتروني.

- ج- اسم خاص بالمستخدم، إشارة & ، اسم الموقع الذي يقدم خدمة البريد الإلكتروني.
د- اسم خاص بالمستخدم، إشارة @ ، اسم الموقع الذي يقدم خدمة البريد الإلكتروني.

14- إن المصطلح WWW يعني:

- أ- الشبكة العنكبوتية العالمية
ب- عنوان الموقع.
ج- محركات البحث
د- متصفح الانترنت.

15- تصنف شبكات الحاسوب حسب المنطقة الجغرافية التي تغطيها إلى:

- أ- شبكات محلية LAN، شبكات واسعة التغطية WAN.
ب- شبكة الخادم والمستخدم Client/Server، الشبكة التناظرية Peer-to-Peer
ج- شبكات محلية LAN، شبكة القيمة المضافة VAN.
د- شبكات واسعة التغطية WAN، شبكة القيمة المضافة VAN.

* الترميز Coding

عمية تحويل إجابات كل سؤال إلى أرقام أو حروف يسهل إدخالها إلى الحاسوب.

مثال: متغير الجنس Sex وهو (ذكر، أنثى)، حيث يعطى: الرقم 1 للذكور و2 للإناث.

○ ذكر. 1

○ أنثى. 2

التخصص: يعطى 1-علمي 2-أدبي 3-معلوماتية 4-مهني كما يلي:

○ علمي. 1

○ أدبي. 2

○ معلوماتية. 3

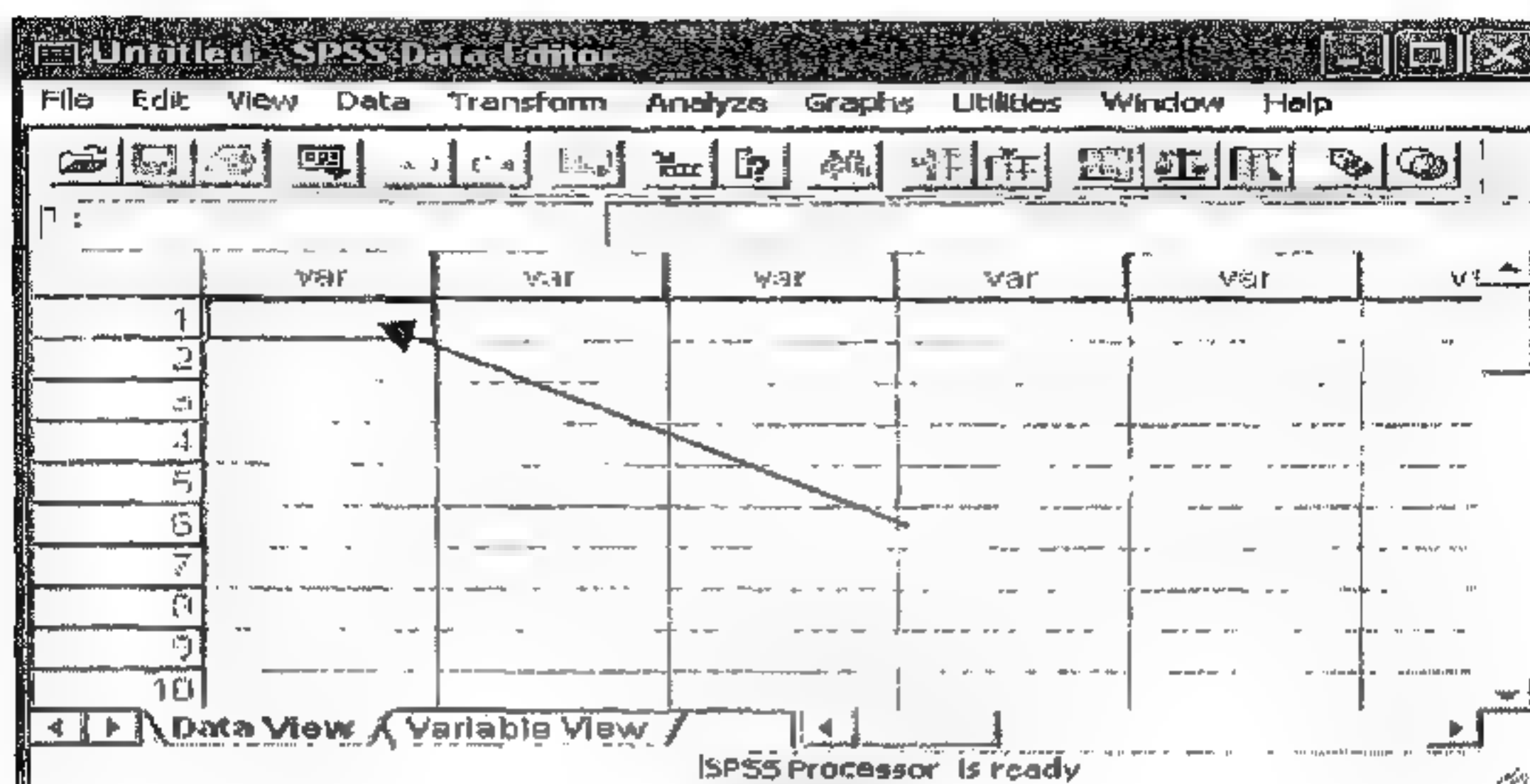
○ مهني. 4

رقم الطالب، الأسئلة:

مثال 1: انشيء ملف جديد اسمه example1

File - New - Data

تظهر لديك الشاشة المبينة أدناه:



لعمل جدول ترميز للمتغيرات الواردة في المثال يتم كما يلي:

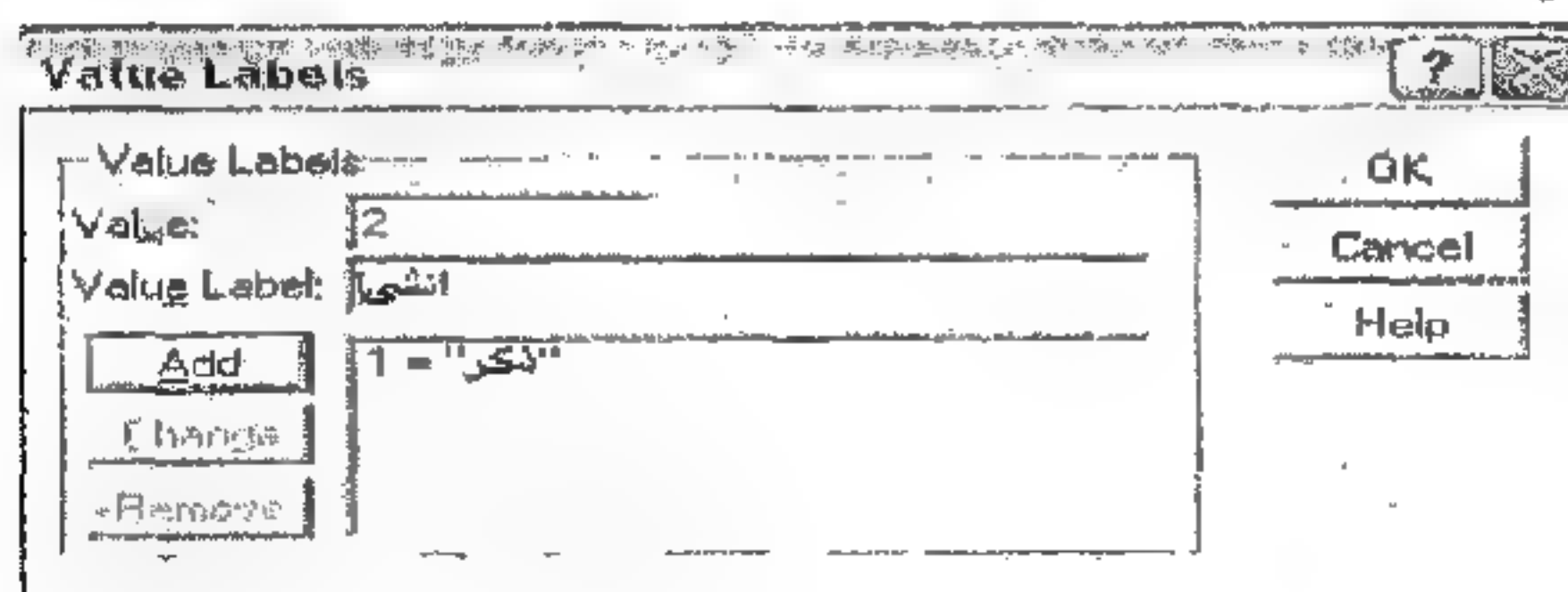
انقر على Variable View الموجودة على شريط الحالة فتظهر الشاشة أدناه

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1										
2										
3										

إدخال متغير رقم الطالب stno

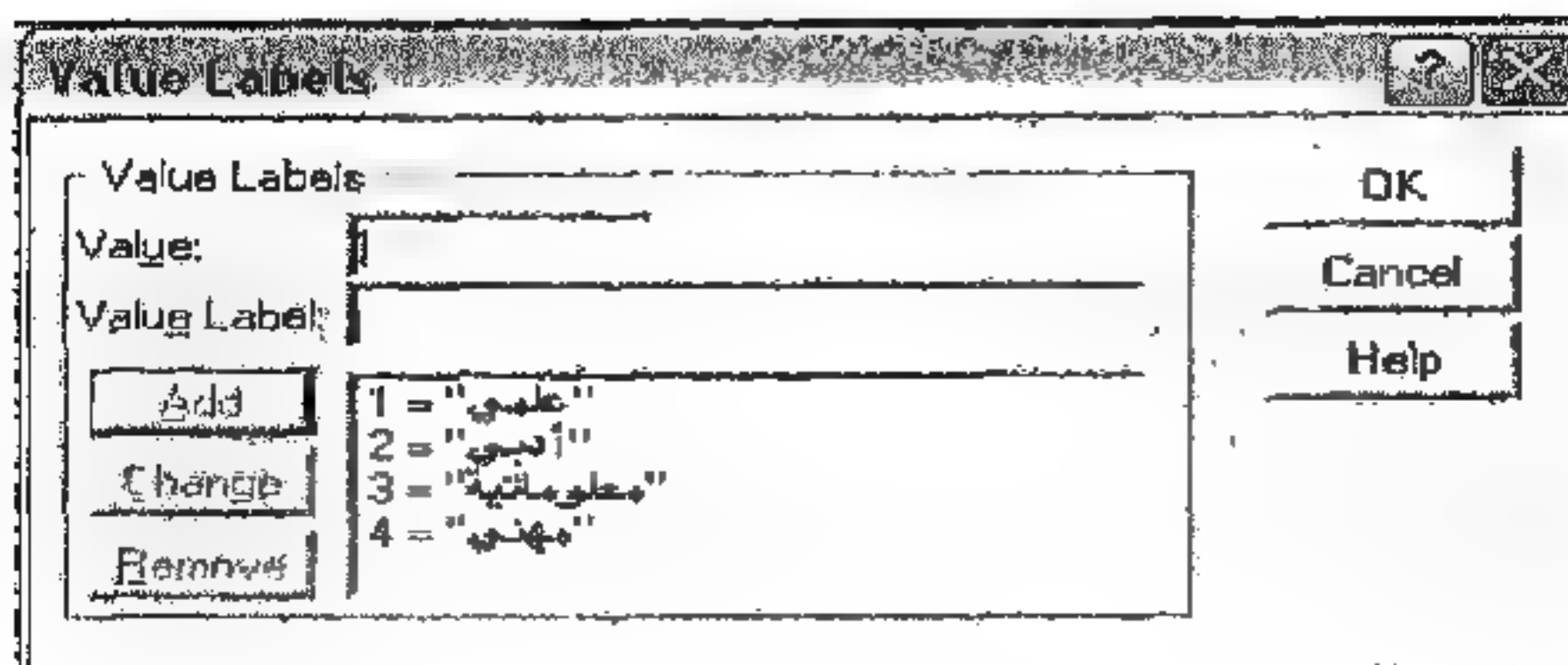
إدخال مستغير الجنس sex وعند وصولنا إلى خانة Values وهي القيم المحتملة للمتغير

نعبأها بالشكل التالي:



إدخال متغير التخصص spec وعند وصولنا إلى خانة Values وهي القيم المحتملة للمتغير

عبأها بالشكل التالي:



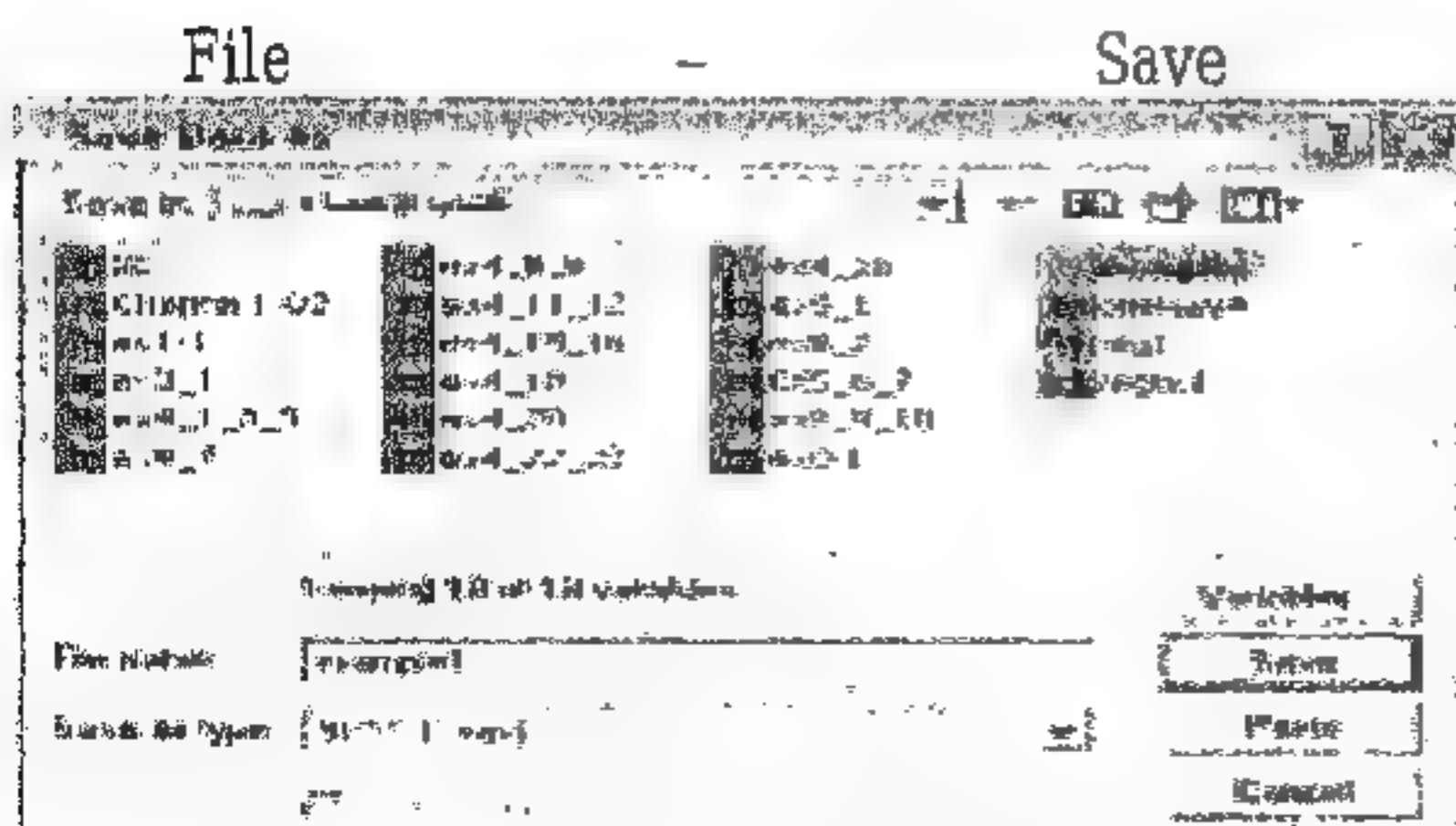
إدخال متغيرات الأسئلة من سؤال 1 إلى سؤال 15 وهي q1 ... q15 وبعد الانتهاء من تعريف المتغيرات تظهر الشاشة كما هو مبين أدناه:

	Name	Type	Width	Decimals	Label	Values	Missing	Column	Align	Measure
1	stno	Numeric	3	0		None	None	3	Center	Scale
2	sex	Numeric	1	0		1 = ذكر, 2 = أنثى	None	8	Right	Scale
3	spec	Numeric	1	0		1 = علمي, 2 = أدبي, 3 = معلوماتية, 4 = مهني	None	8	Right	Scale
4	q1	Numeric	1	0		None	None	3	Center	Scale
5	q2	Numeric	1	0		None	None	3	Center	Scale
6	q3	Numeric	1	0		None	None	3	Center	Scale
7	q4	Numeric	1	0		None	None	3	Center	Scale
8	q5	Numeric	1	0		None	None	3	Center	Scale
9	q6	Numeric	1	0		None	None	3	Center	Scale
10	q7	Numeric	1	0		None	None	3	Center	Scale
11	q8	Numeric	1	0		None	None	3	Center	Scale
12	q9	Numeric	1	0		None	None	3	Center	Scale
13	q10	Numeric	1	0		None	None	3	Center	Scale
14	q11	Numeric	1	0		None	None	3	Center	Scale
15	q12	Numeric	1	0		None	None	3	Center	Scale
16	q13	Numeric	1	0		None	None	3	Center	Scale
17	q14	Numeric	1	0		None	None	3	Center	Scale
18	q15	Numeric	1	0		None	None	3	Center	Scale

إدخال البيانات Input Data

	slno	sex	age	q1	q2	q3	q4	q5	q6	q7	q8	q9	q10	q11	q12	q13	q14	q15
1	1	1	3	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1
2	2	1	3	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
3	3	1	3	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1
4	4	1	3	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1
5	5	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1
6	6	1	1	1	1	0	1	1	0	1	1	1	1	1	0	1	1	1
7	7	1	1	1	1	0	0	1	0	1	1	1	0	1	0	1	1	1
8	8	1	1	1	0	0	0	1	0	1	1	1	0	1	1	1	1	1
9	9	1	1	1	0	0	1	1	0	1	1	1	1	0	1	1	1	1
10	10	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	1
11	11	2	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
12	12	2	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
13	13	2	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1
14	14	2	1	0	1	0	0	0	0	0	1	1	0	0	0	1	1	1
15	15	2	1	0	0	1	0	0	0	1	1	1	1	1	1	1	1	1
16	16	2	2	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1
17	17	2	2	1	1	0	1	1	0	0	1	1	1	1	1	1	1	1
18	18	2	2	0	0	0	1	0	1	1	1	1	1	1	0	0	1	0
19	19	2	2	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1
20	20	2	2	1	1	0	0	1	0	1	1	1	1	0	0	0	1	1
21	21	2	2	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1
22	22	2	2	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
23	23	2	2	1	0	0	0	0	0	0	0	1	1	0	1	1	1	1
24	24	2	2	0	0	0	1	0	1	0	0	1	0	0	1	0	1	0
26	25	2	3	1	0	0	0	1	0	0	1	1	0	0	0	1	0	1
26	26	2	3	1	0	0	1	1	0	1	1	1	1	0	1	1	1	1
27	27	2	3	1	1	0	1	1	0	1	0	1	1	1	0	1	1	1
28	28	2	3	1	1	0	1	0	0	1	0	1	1	1	0	1	1	1
29	29	2	3	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1
30	30	2	3	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1

حفظ البيانات



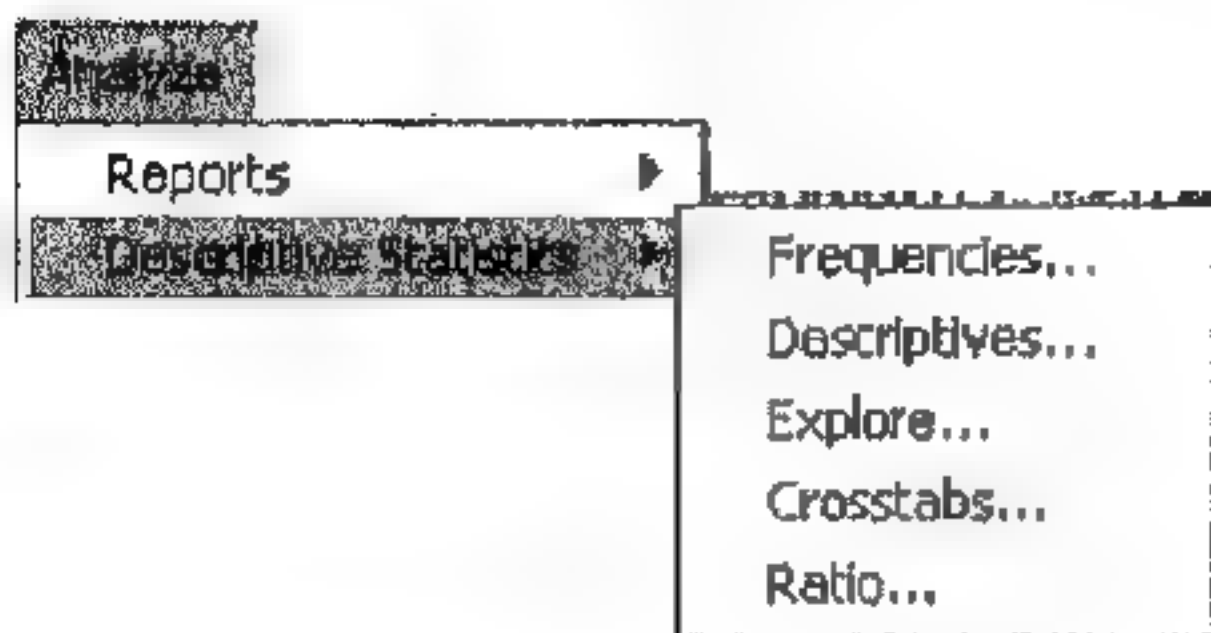
نكتب اسم الملف example1 ثم نقر زر Save فيحفظ الملف بهذا الاسم.

التحليلات المطلوبة:

1- الإحصاء الوصفي Descriptive Statistics

- Descriptive Statistics - Frequencies Analyze -

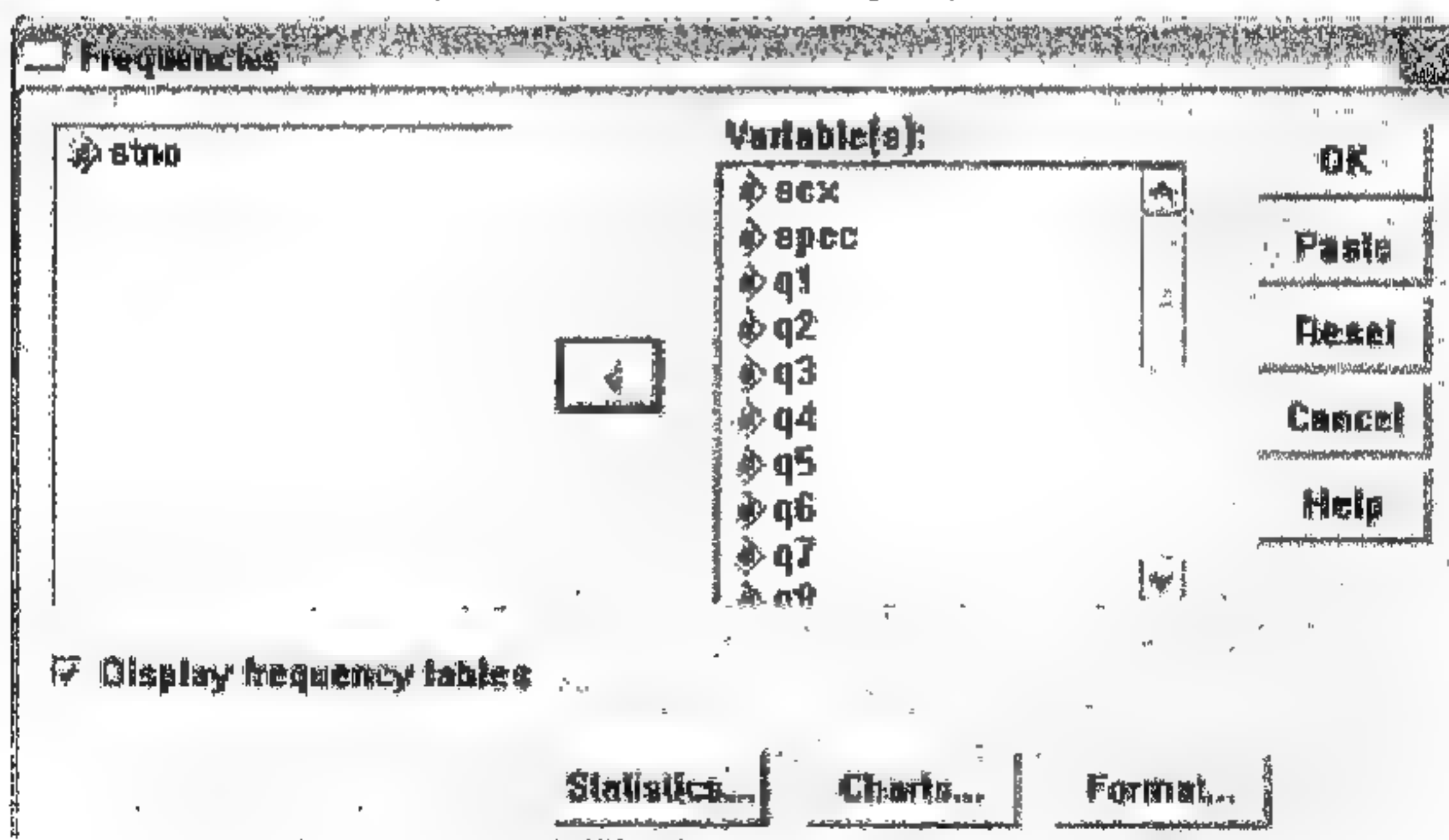
Frequencies: وصف توزيع أفراد العينة حسب أحد المتغيرات من النوع الاسمي.



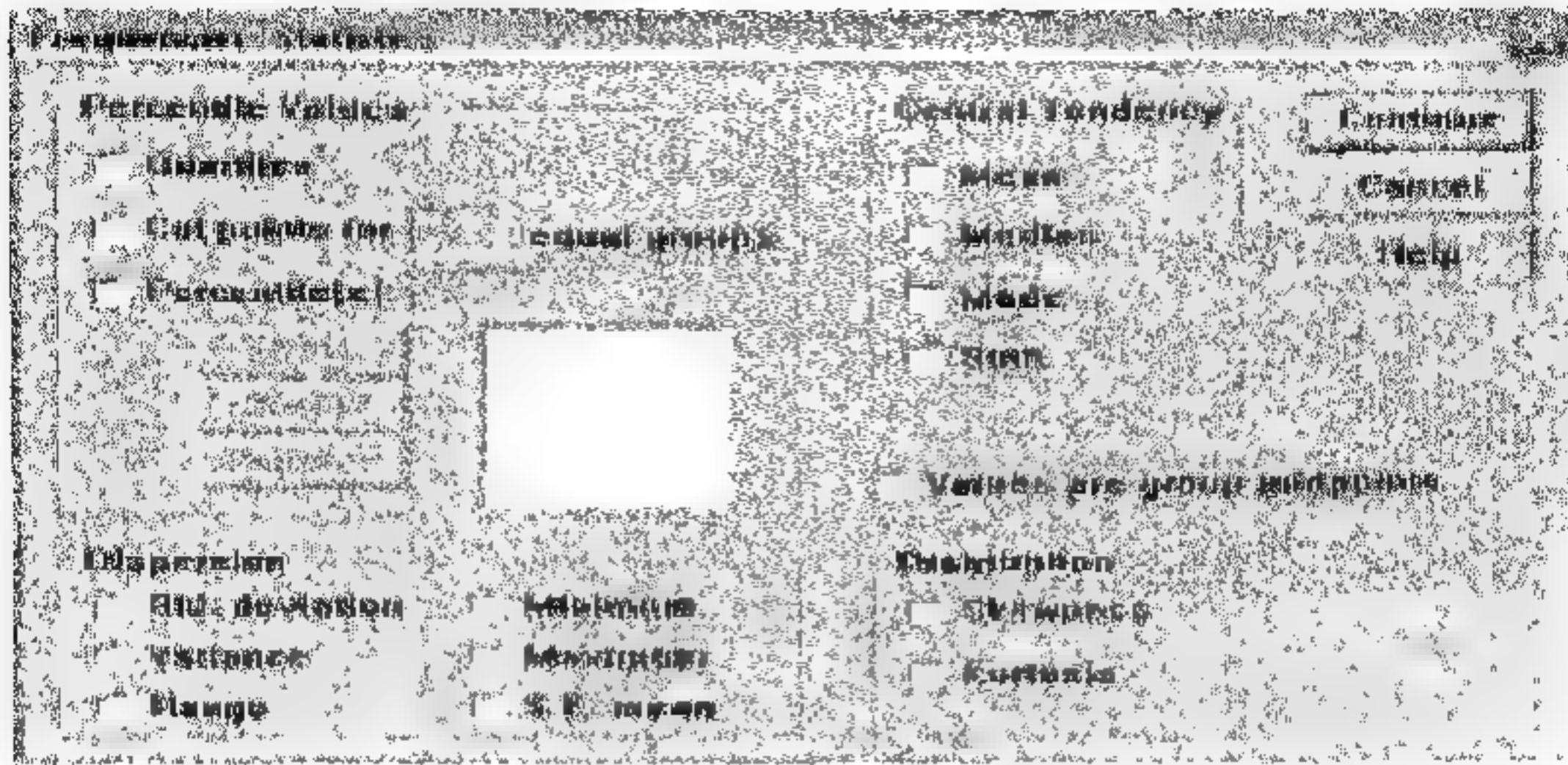
بعض أسئلة الدراسة:

س1: ما نسبة الذكور والإناث في عينة الدراسة؟ (متغير الجنس sex)

س2: ما عدد أفراد العينة في كل فئة من فئات التخصص؟ (متغير التخصص spec)



بالضغط على زر Statistics... (الاحصاءات) تظهر الشاشة المبينة أدناه:



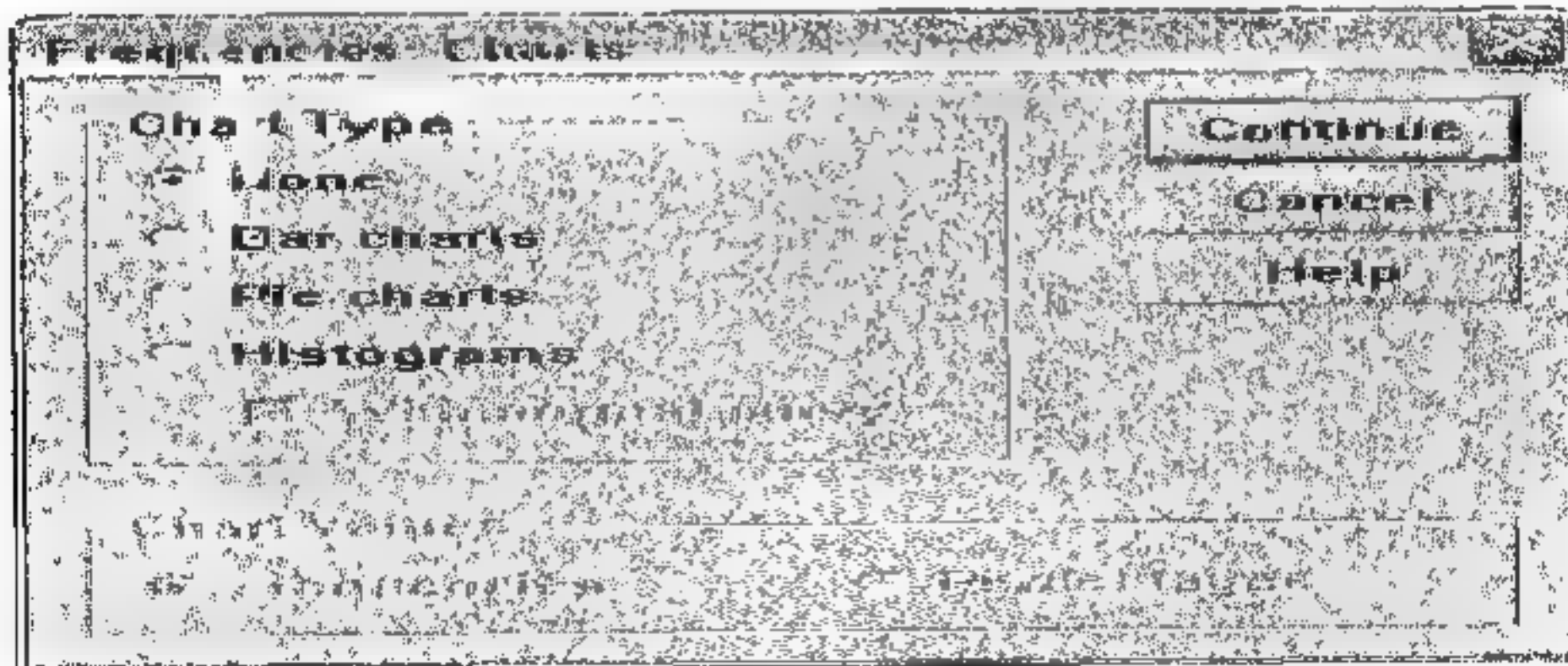
يستخدم لاستخراج بعض إحصاءات مقاييس التزعة المركزية Central Tendency: مثل Mean الوسط الحسابي، Median الوسيط، Mode المنوال، Sum المجموع.

يستخدم لاستخراج بعض إحصاءات مقاييس التشتت: مثل Std. deviation الانحراف المعياري، Variance التباين، Range المدى.

يستخدم لاستخراج بعض إحصاءات القيم المئينية: مثل Quartiles الربعيات، Percentiles المئينات.

يستخدم لاستخراج بعض إحصاءات شكل التوزيع: مثل Skewness الالتواء، Kurtosis التفرطح.

بالضغط على زر Charts... (الرسومات البيانية) تظهر الشاشة المبينة أدناه:



نوع الرسم البياني: Chart Type

1- الأعمدة Bar charts

2- الدوائر Pie charts

3- الرسم البياني مع المنحنى الطبيعي Histogram

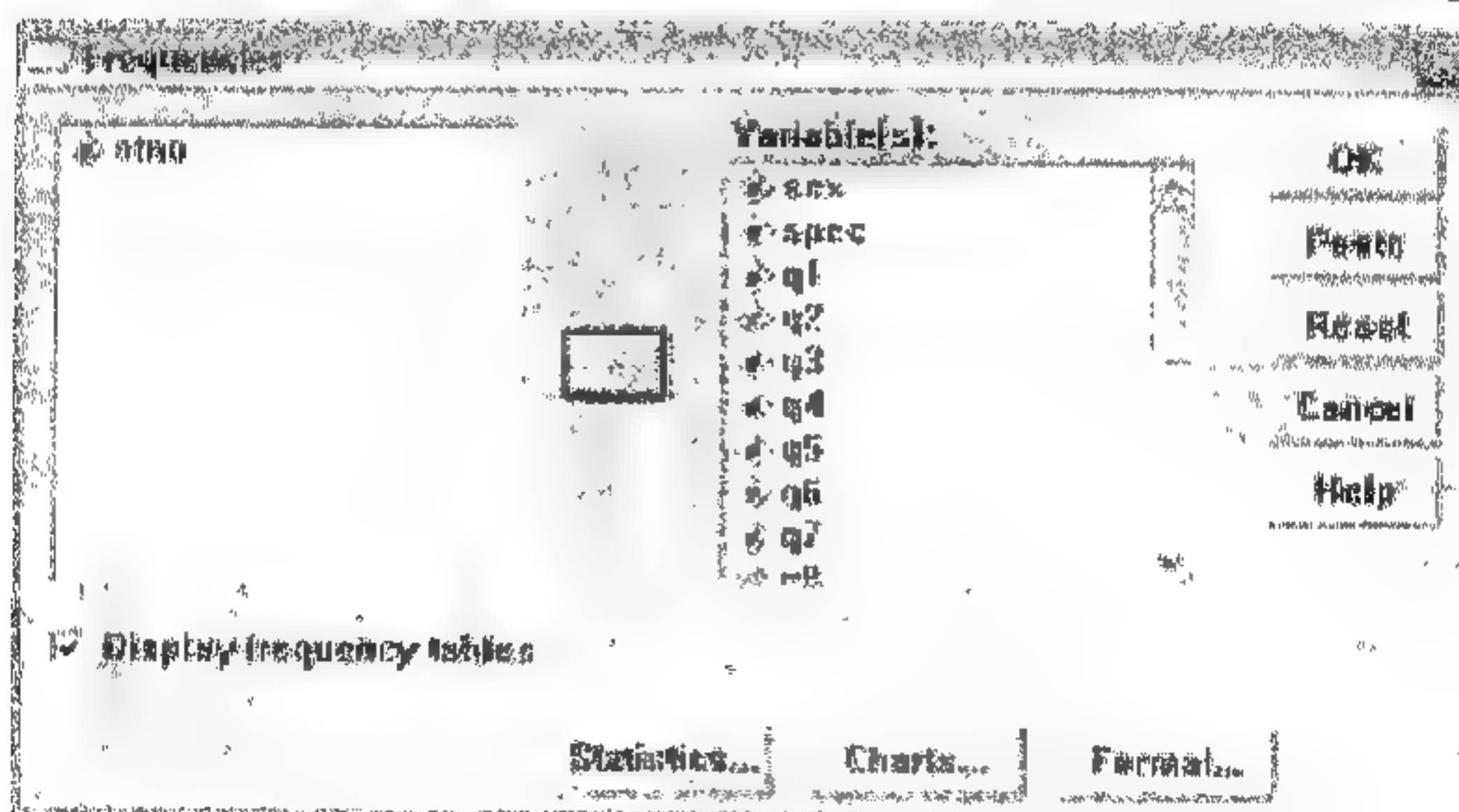
* لإيجاد الوسط Mean، الوسيط Median، المنوال Mode، المجموع Sum

Analyze - Descriptive Statistics - Frequencies...



نعمل على اختيار المتغيرات وهي sex, spec, q1...q15 ثم نضغط زر Statistics ونحدد

ما نريد



نحدد Mean, Median, Mode, Sum ثم نضغط زر Continue

Frequency Statistics

Percentile Values

☒ Quantiles

☐ Cut points for equal groups

☒ Percentiles

50

Central Tendency

☒ Mean

☒ Median

☒ Mode

☒ Sum

Values are group midpoints

Dispersion

☒ Std. deviation

☒ Minimum

☒ Variance

☒ Maximum

☒ Range

☒ S.E. mean

Distribution

☒ Skewness

☒ Kurtosis

Continue

Cancel

Help

تظهر المخرجات المبينة أدناه:

لاحصاءات المطلوبة لكل متغير من متغيرات الدراسة:

Table 1

	q1	q2	q3	q4	q5	q6	q7	q8	q9	q10
Mean	50	50	50	50	50	50	50	50	50	50
Std. deviation	1.41	1.41	1.41	1.41	1.41	1.41	1.41	1.41	1.41	1.41
Median	50	50	50	50	50	50	50	50	50	50
Mode	50	50	50	50	50	50	50	50	50	50
Std. deviation	1.41	1.41	1.41	1.41	1.41	1.41	1.41	1.41	1.41	1.41
Skewness	-.318	-.318	-.318	-.318	-.318	-.318	-.318	-.318	-.318	-.318
Kurtosis	1.954	1.954	1.954	1.954	1.954	1.954	1.954	1.954	1.954	1.954
Range	1	1	1	1	1	1	1	1	1	1
Minimum	49	49	49	49	49	49	49	49	49	49
Maximum	51	51	51	51	51	51	51	51	51	51
Sum	500	500	500	500	500	500	500	500	500	500
Percentile 1	49	49	49	49	49	49	49	49	49	49
Percentile 2	50	50	50	50	50	50	50	50	50	50
Percentile 3	51	51	51	51	51	51	51	51	51	51

توزيع أفراد العينة حسب متغيري الجنس sex، التخصص spec، الاسئلة q1...q15

Frequency Table

SEX

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid ذكر	10	33.3	33.3	33.3
Valid أنثى	20	66.7	66.7	100.0
Total	30	100.0	100.0	

SPEC

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid علمي	11	36.7	36.7	36.7
Valid أدبي	9	30.0	30.0	66.7
Valid معلوماتية	10	33.3	33.3	100.0
Total	30	100.0	100.0	

Q1

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0	5	16.7	16.7	16.7
Valid 1	25	83.3	83.3	100.0
Total	30	100.0	100.0	

Q2

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0	10	33.3	33.3	33.3
Valid 1	20	66.7	66.7	100.0
Total	30	100.0	100.0	

Q3

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0	25	83.3	83.3	83.3
Valid 1	5	16.7	16.7	100.0
Total	30	100.0	100.0	

Q4

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0	14	46.7	46.7	46.7
Valid 1	16	53.3	53.3	100.0
Total	30	100.0	100.0	

Q5

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0	7	23.3	23.3	23.3
Valid 1	23	76.7	76.7	100.0
Total	30	100.0	100.0	

Q6

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0	18	60.0	60.0	60.0
Valid 1	12	40.0	40.0	100.0
Total	30	100.0	100.0	

Q7

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0	5	16.7	16.7	16.7
1	25	83.3	83.3	100.0
Total	30	100.0	100.0	

Q8

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0	5	16.7	16.7	16.7
1	25	83.3	83.3	100.0
Total	30	100.0	100.0	

Q9

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0	1	3.3	3.3	3.3
1	29	96.7	96.7	100.0
Total	30	100.0	100.0	

Q10

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0	6	20.0	20.0	20.0
1	24	80.0	80.0	100.0
Total	30	100.0	100.0	

Q11

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0	10	33.3	33.3	33.3
1	20	66.7	66.7	100.0
Total	30	100.0	100.0	

Q12

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0	10	33.3	33.3	33.3
1	20	66.7	66.7	100.0
Total	30	100.0	100.0	

Q13

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0	4	13.3	13.3	13.3
1	26	86.7	86.7	100.0
Total	30	100.0	100.0	

Q14

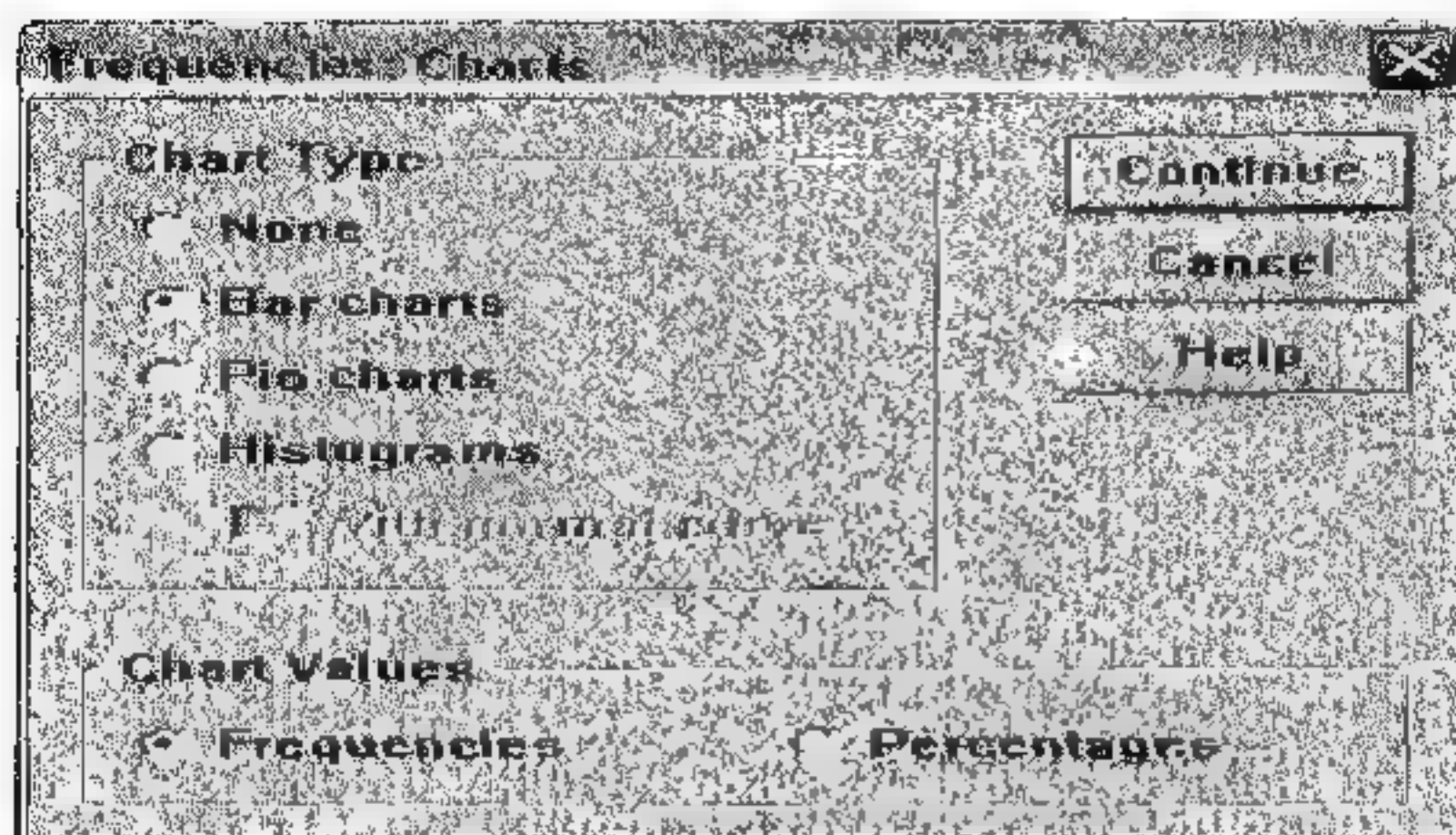
	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0	2	6.7	6.7	6.7
1	28	93.3	93.3	100.0
Total	30	100.0	100.0	

Q15

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0	2	6.7	6.7	6.7
1	28	93.3	93.3	100.0
Total	30	100.0	100.0	

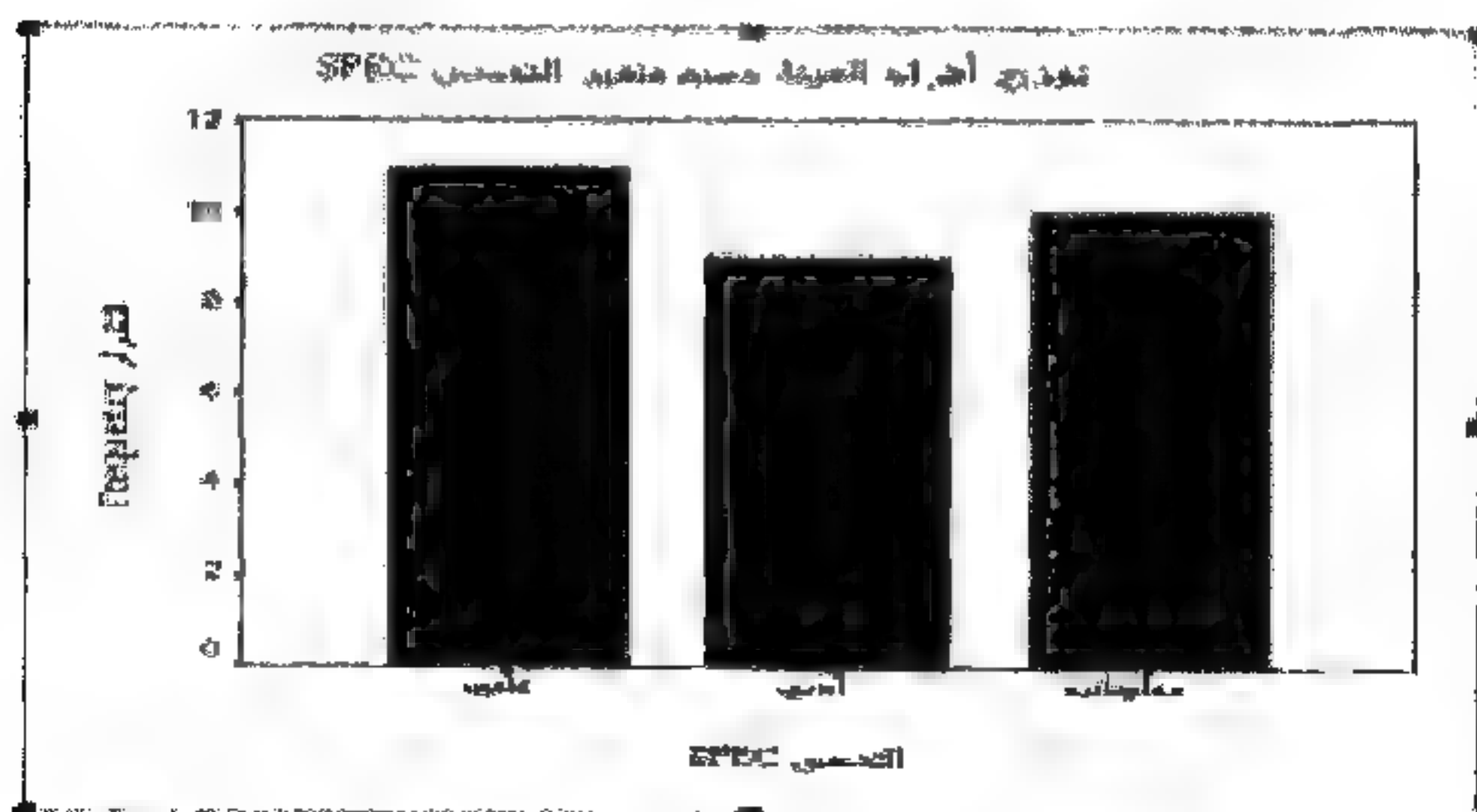
تمثيل النتائج باستخدام الرسم البياني

بالضغط على زر Charts... تظهر الشاشة المبينة أدناه:



لنختار نوع Bar charts ثم نضغط زر Continue ثم نضغط زر Ok

Bar Chart



النتائج:

تتكون العينة من (30) فرداً، عدد الذكور (10) طالباً، وعدد الاناث (20) طالبة.
تتكون العينة من (30) فرداً، الفرع العلمي (11)، الفرع الأدبي (9)، فرع المعوماتية (10)

Frequency Table

SEX

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid ذكر	10	33.3	33.3	33.3
انثى	20	66.7	66.7	100.0
Total	30	100.0	100.0	

SPEC

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid علمي	11	36.7	36.7	36.7
أدبي	9	30.0	30.0	66.7
معلوماتية	10	33.3	33.3	100.0
Total	30	100.0	100.0	

بالنسبة للسؤال الأول، اجاب عليه (25) اجابة صحيحة، واجاب عليه (5) اجابة

خاطئة.

بالنسبة للسؤال الثاني، اجاب عليه (20) اجابة صحيحة، واجاب عليه (10) اجابة

خاطئة.

Q1

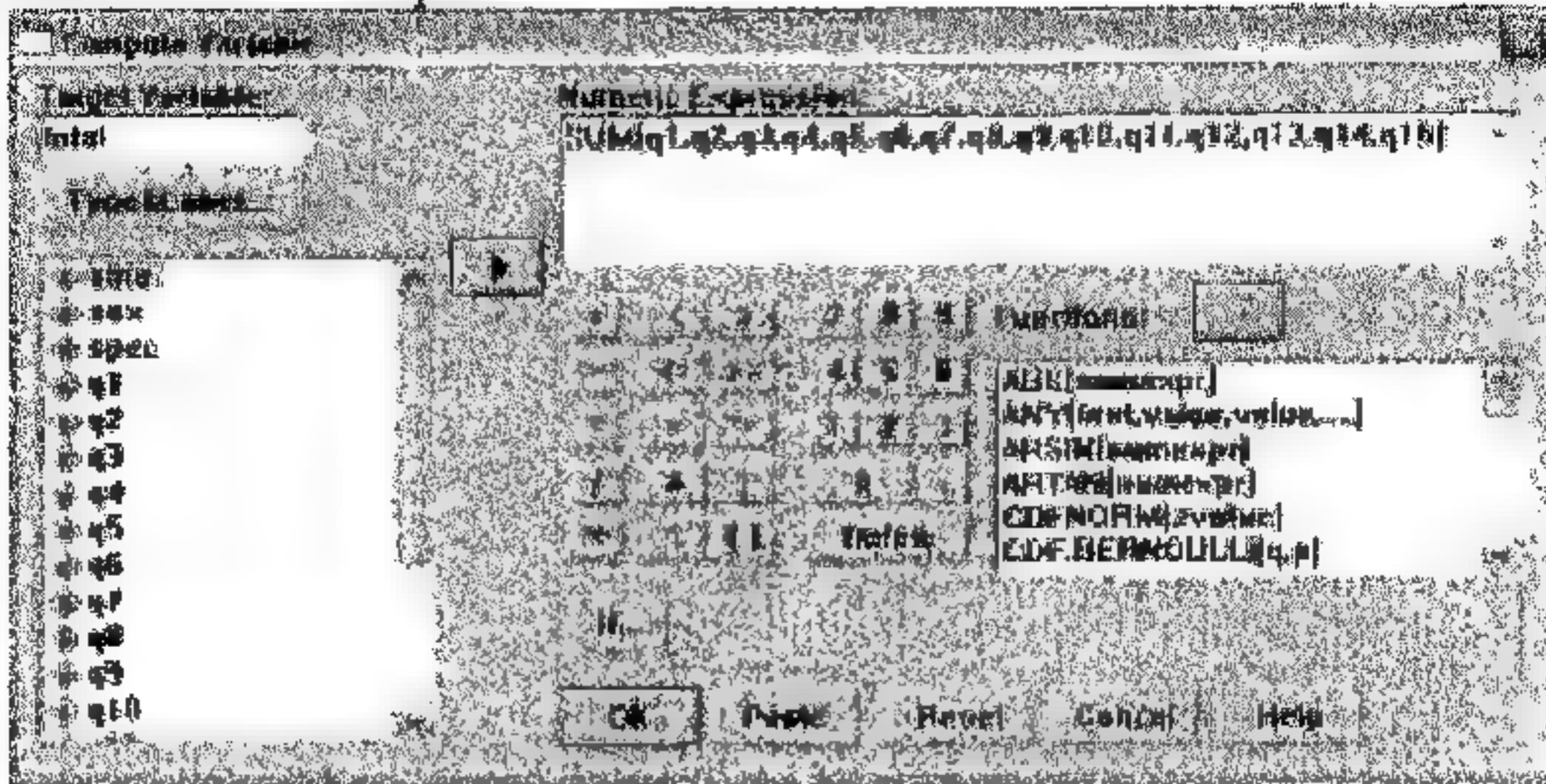
	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0	5	16.7	16.7	16.7
1	25	83.3	83.3	100.0
Total	30	100.0	100.0	

Q2

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 0	10	33.3	33.3	33.3
1	20	66.7	66.7	100.0
Total	30	100.0	100.0	

* لإيجاد متغير اسمه total وهو مجموع علامات كل طالب على الاختبار

Transform – Compute...

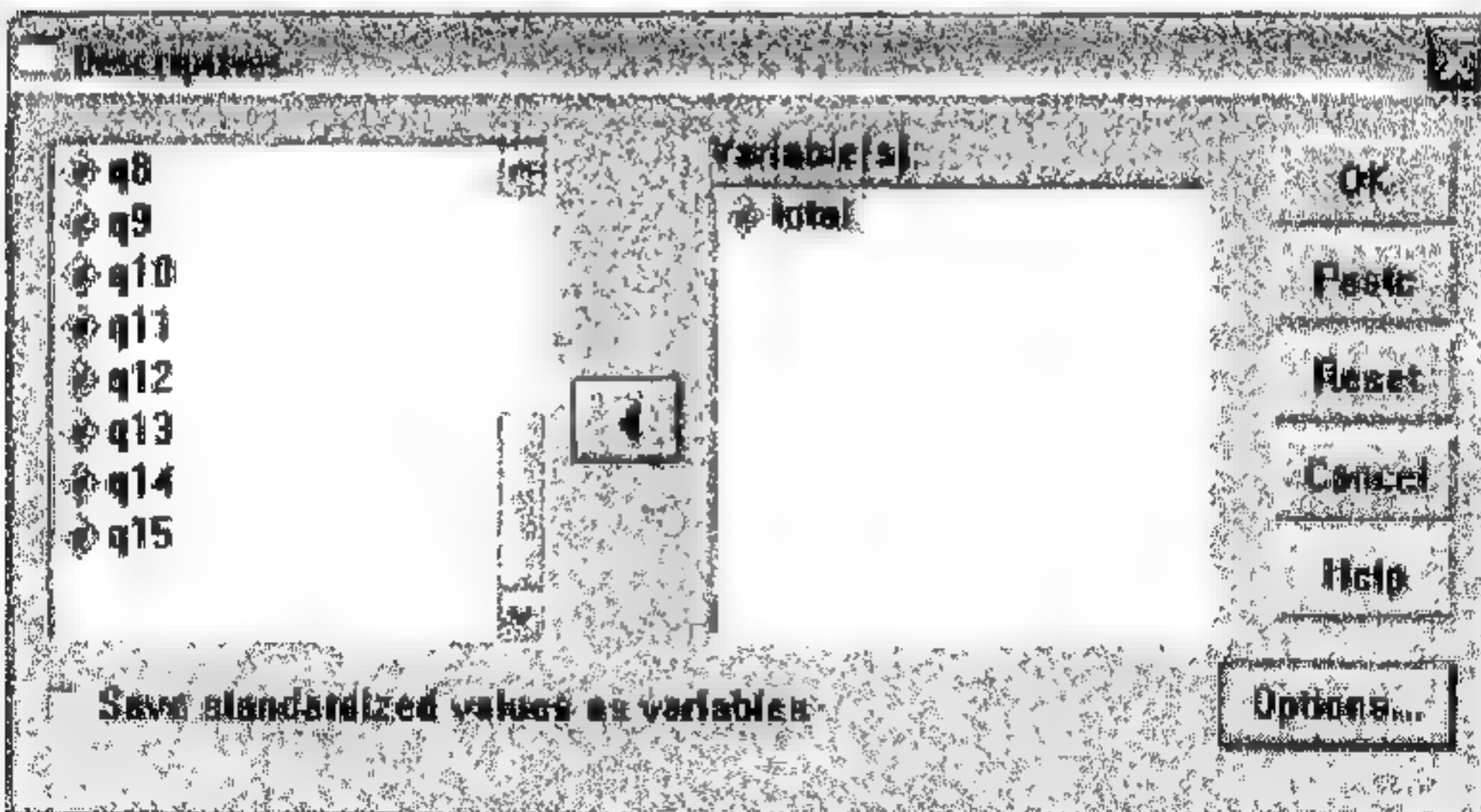


يظهر متغير جديد اسمه total وهو عبارة عن مجموع علامة كل طالب في جميع الاسئلة

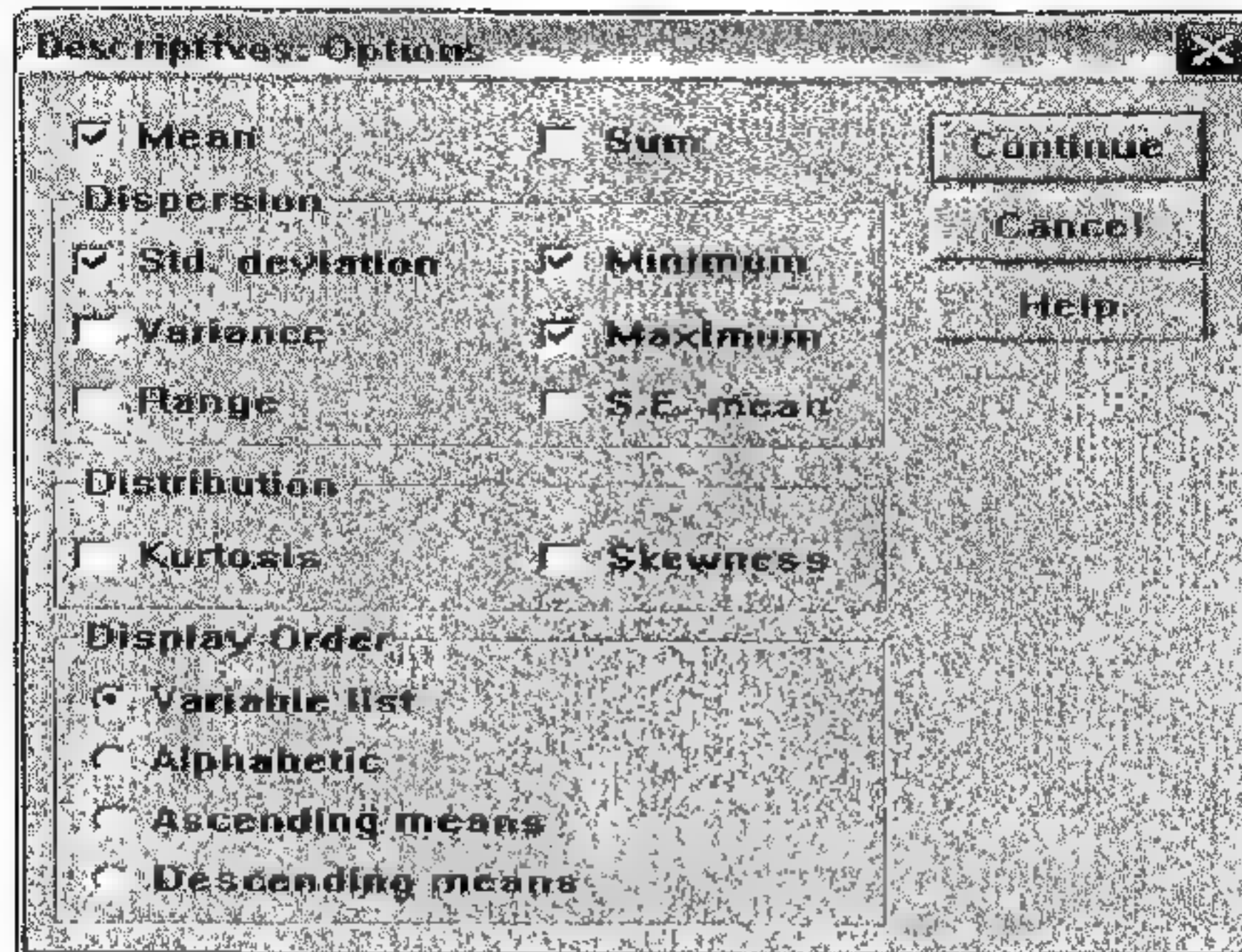
من 1- 15 باستخدام الاقتران Sum من قائمة Functions:

* لإيجاد الوسط Mean، الوسيط Median، المنوال Mode، المجموع Sum

Analyze – Descriptive Statistics – Descriptives...



اضغط زر Options... تظهر لديك الشاشة المبينة أدناه:



اختر الاحصاءات المطلوبة ثم اضغط زر Continue ثم Ok تظهر النتائج أدناه:

→ Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
TOTAL	30	3.00	14.00	10.8667	3.00268
Valid N (listwise)	30				

النتيجة:

تراوحت علامات الطلاب بين 3 إلى 14، وبلغ المتوسط الحسابي لها 10.8667 بانحراف معياري 3.00268، وكان عدد العينة 30 شخصاً.

* يستخدم الاجراء الاحصائي Explore للتحقق من الخطوة الأساسية قبل إجراء التحليلات الاحصائية وهي فحص البيانات ومحاولة تصحيح الاخطاء إن وجدت مثل الارقام غير المنطقية أو الشذاة أو التحقق من أن توزيع المتغير طبيعياً، أو التحقق من شرط تجانس التباين.

* يستخدم إجراء Explore لعمل ما يلي:

حساب الاحصاءات الوصفية.

عمل بعض الرسومات لتوضيح شكل توزيع المتغيرات مثل Histograms, Stem-and-

Leaf Plot, Box Plot

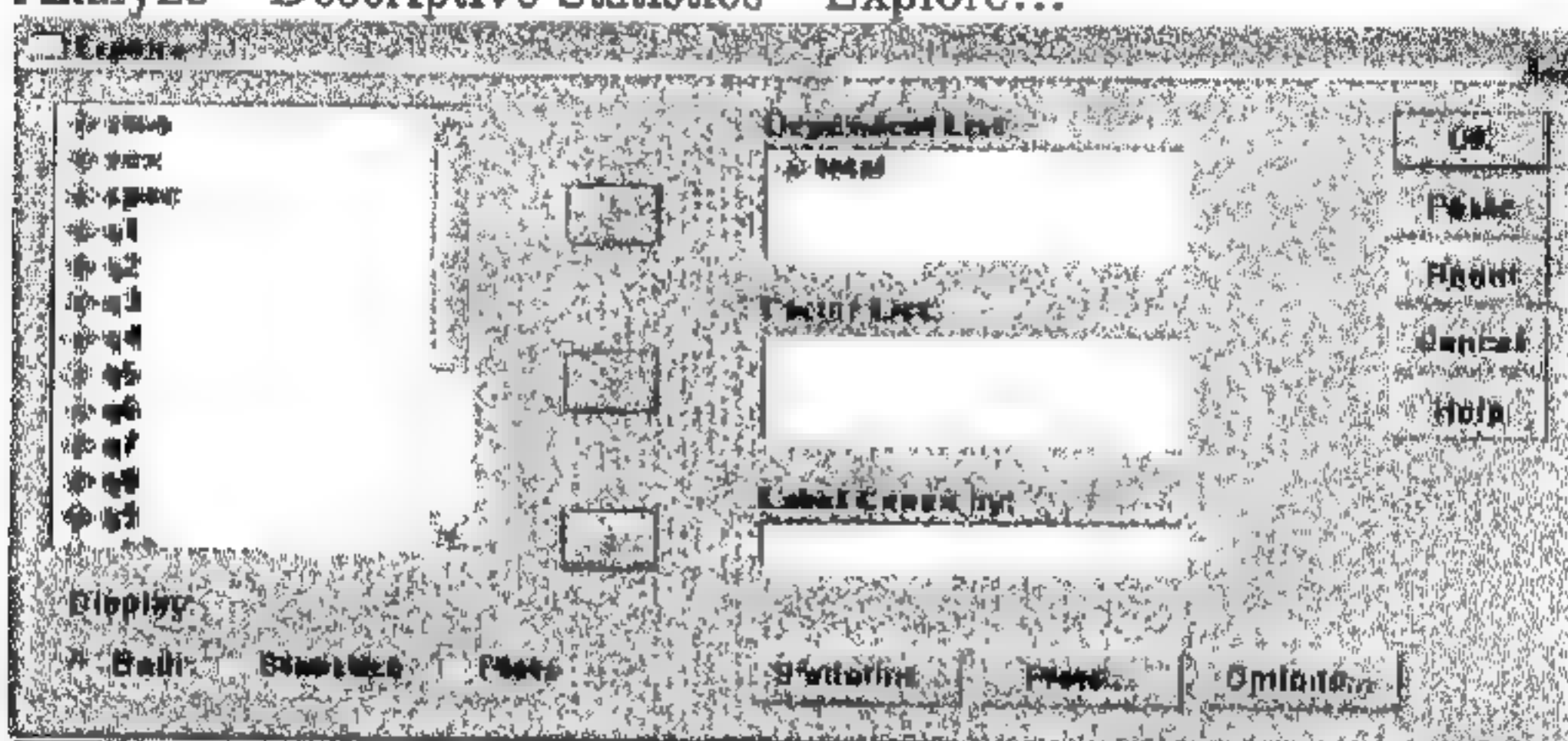
اختبار التوزيع الطبيعي عن طريق اختبار Shapiro Wilks واختبار Lilliefors

اختبار تجانس التباين Homogeneity of Variances عن طريق اختبار

Levene-Test

* حساب الاحصاءات الوصفية للمتغير total

Analyze – Descriptive Statistics – Explore...



Descriptives			
		Statistic	Std. Error
TOTAL	Mean	10.8667	54821
	95% Confidence Interval for Mean	9.7454	
	Lower Bound	11.9879	
	Upper Bound	11.0928	
	5% Trimmed Mean	11.0928	
	Median	12.0000	
	Variance	9.016	
	Std. Deviation	3.00268	
	Minimum	3.00	
	Maximum	14.00	
	Range	11.00	
	Interquartile Range	3.2500	
	Skewness	-1.050	.427
	Kurtosis	.384	.833

* استخدام اختبار CHI-Square

الجدول المتقاطع-مربع كاي Crosstabulation Table: Chi-Square

أن مربع كاي χ^2 من الإحصائيات الهامة ولها عدة استخدامات منها الكشف عن عملية الاستقلالية Independence بين متغيرين عندما تكون هناك تكرارات ويكون لكل متغير عدة مستويات محدودة.

خطوات عمل الجدول المتقاطع-مربع كاي

مثال: البيانات التالية تمثل الأطوال للآباء (Fathers) ويمثل ثلاث مستويات (طويل، متوسط، قصير) لمجموعة من الآباء والأطوال لمجموعة من الأبناء (Suns) ويمثل ثلاث مستويات (طويل، متوسط، قصير) والتي تشمل 20 أب مع أبنائهم.

المطلوب: هل يوجد ارتباط بين المتغيرين باستخدام مربع كاي χ^2 .

1- تعريف المتغيرات وهي:

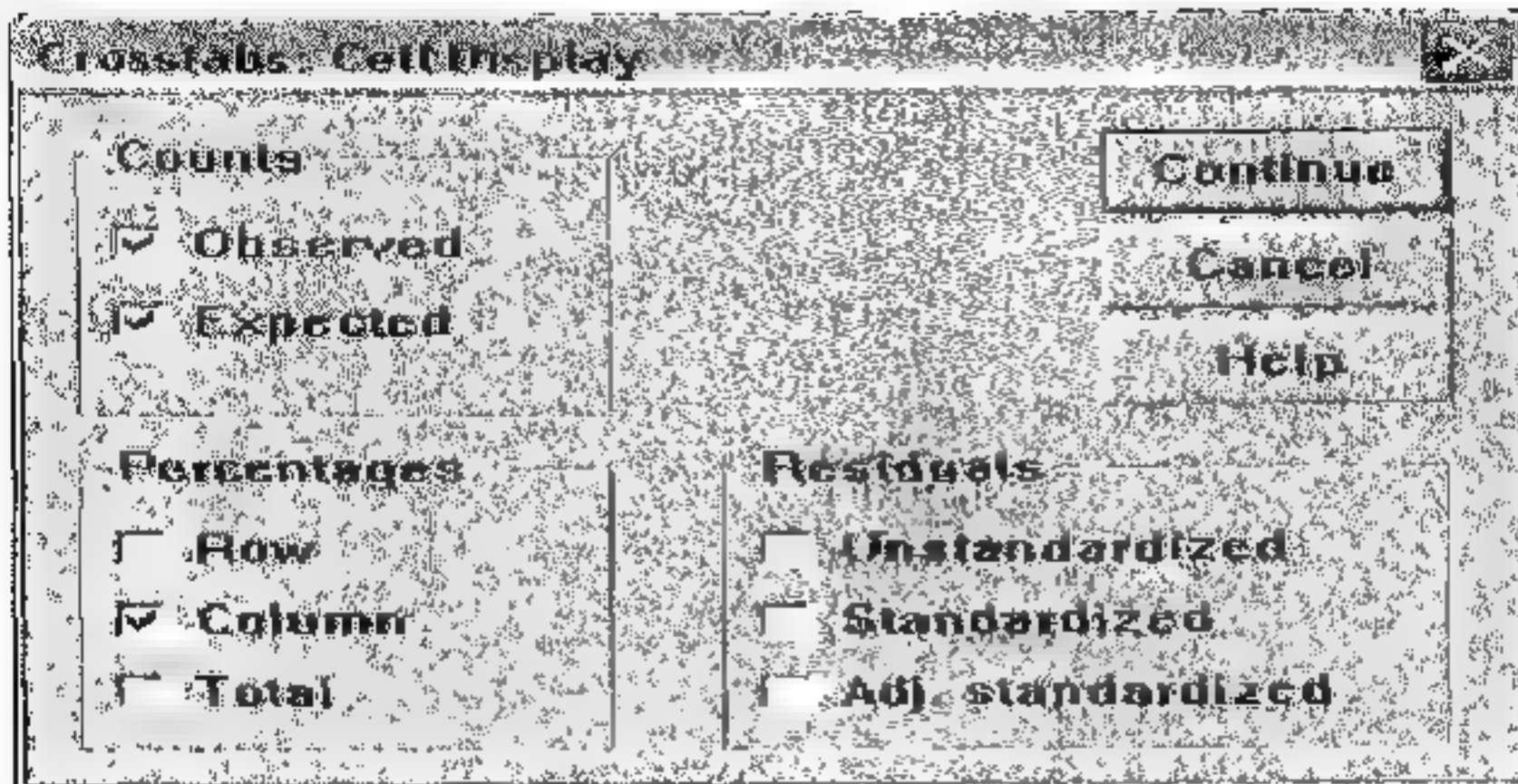
Father ولها ثلاث مستويات: (Long , Medium , Short)

Suns ولها ثلاث مستويات: (Long , Medium , Short)

Name	Type	Width	Decimals	Label	Values	Minimum	Maximum	Range	Center	Scale
FATHER	Numeric	8	0		None	None	8		Center	Scale
SUNS	Numeric	8	0		None	None	8		Center	Scale

Value Label	Label
1	Long
2	Medium
3	Short

Analyze - Descriptive Statistics - Crosstabs...



من الخيار Counts نحدد التكرار المشاهد Observed والتكرار المتوقع Expected
 من الخيار Percentage نختار العمود Columns ويعطي النسبة المئوية لمتغير مستقل
 وبعدها نضغط على زر الاستمرار Continue ثم نضغط زر الموافقة OK
 تظهر شاشة المخرجات أدناه:

Crosstabs

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
\$ UNS * FATHERS	20	100.0%	0	.0%	20	100.0%

\$ UNS * FATHERS Crosstabulation

			FATHERS			Total
			Long	Medium	Short	
\$ UNS	Long	Count	5	2	0	7
		Expected Count	2.1	2.8	2.1	7.0
		% within FATHERS	83.3%	25.0%	0%	35.0%
	Medium	Count	1	6	0	7
		Expected Count	2.1	2.8	2.1	7.0
		% within FATHERS	16.7%	75.0%	0%	35.0%
	Short	Count	0	0	6	6
		Expected Count	1.8	2.4	1.8	6.0
		% within FATHERS	0%	0%	100.0%	30.0%
Total	Count	6	8	6	20	
	Expected Count	6.0	8.0	6.0	20.0	
	% within FATHERS	100.0%	100.0%	100.0%	100.0%	

يمثل نتائج التكرارات ونسبها في الخلايا والتي تم اضافتها في خيار Cell فقد تم اضافة
 expected وكذلك percentage .

Chi Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	20.667 ^a	4	.000
Likelihood Ratio	29.439	4	.000
Linear-by-Linear Association	14.784	1	.000
N of Valid Cases	20		

a. 0 cells (100.0%) have expected count less than 5. The minimum expected count is 1.80

السطر الأول من الجدول هو المطلوب والمهم حيث أن قيمة مربع كاي - Pearson Chi-Square تساوي (26.667) وهذا يعني أن هناك علاقة قوية وموجبة بين أطوال الآباء والأبناء حيث أن مستوى المعنوية يساوي 0.000 .

Symmetric Measures

		Value	Asymp. Std. Error ^a	Approx. T ^b	Approx. Sig.
Internal by Internal	Pearson's R	.882	.067	7.257	.000 ^c
Ordinal by Ordinal	Spearman Correlation	.878	.083	7.796	.000 ^c
N of Valid Cases		20			

a. Not assuming the null hypothesis.

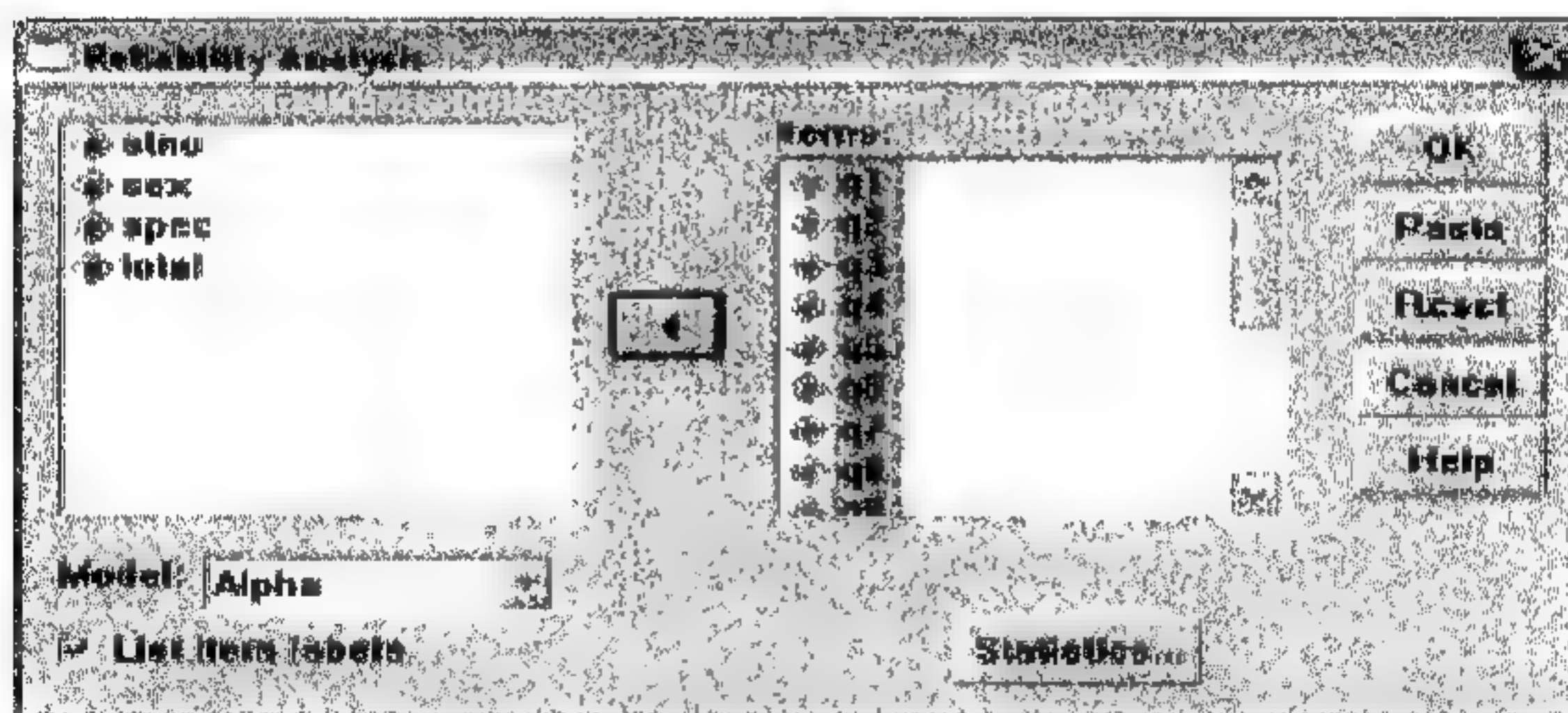
b. Using the asymptotic standard error assuming the null hypothesis.

c. Based on normal approximation.

* الثبات Reliability : اختبار كرونباخ الفا Cronbach Alpha

يستخدم هذا الاختبار لتحديد ثبات الاختبار

Analyze - Scale - Reliability Analysis...



نختار المتغيرات المطلوبة وهي الاسئلة من q1-q15 ونضعها في قائمة Items، ثم نضغط زر Ok، فتظهر شاشة المخرجات أدناه:

Reliability

```

***** Method 1 (spss server) will be used for this analysis *****
R E L I A B I L I T Y   A N A L Y S I S -   S C A L E   (A L P H A)
1.      Q1
2.      Q2
3.      Q3
4.      Q4
5.      Q5
6.      Q6
7.      Q7
8.      Q8
9.      Q9
10.     Q10
11.     Q11
12.     Q12
13.     Q13
14.     Q14
15.     Q15
Reliability Coefficient
N of Cases =      30.0
N of Items = 15
Alpha =      .7860
    
```

وتكون اصغر قيمة مقبولة لمعامل كرونباخ الفا هي 0.6 وأفضل قيمة مقبولة هي (0.7-0.8) وكلما زادت القيمة كانت افضل.
في هذا التحليل قيمة الفا = 0.786 وهي جيدة.

الملاحق

ملحق 1: جداول التوزيعات الاحتمالية.

جدول التوزيع الطبيعي. Z

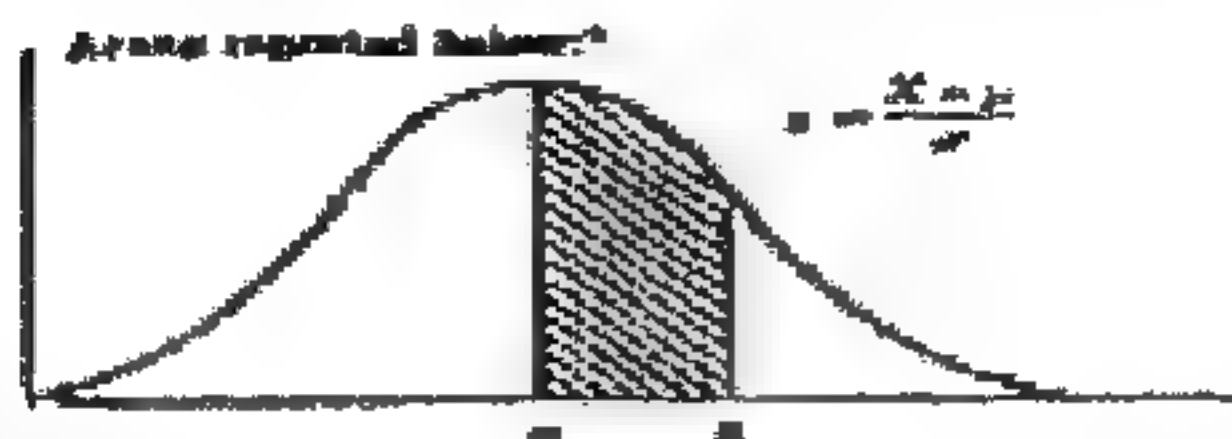
جدول توزيع ت. t

جدول توزيع ف. F

جدول توزيع كا². χ^2

جدول التوزيع الطبيعي. Z

الجدول أدناه يعطي المساحة تحت منحنى التوزيع الطبيعي القياسي (المساحة ما بين المتوسط وقيمة Z)



Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1025	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1913	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2518	.2549
0.7	.2580	.2612	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4014
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4163	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4983	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987									
3.1	.4997									
3.2	.4999									

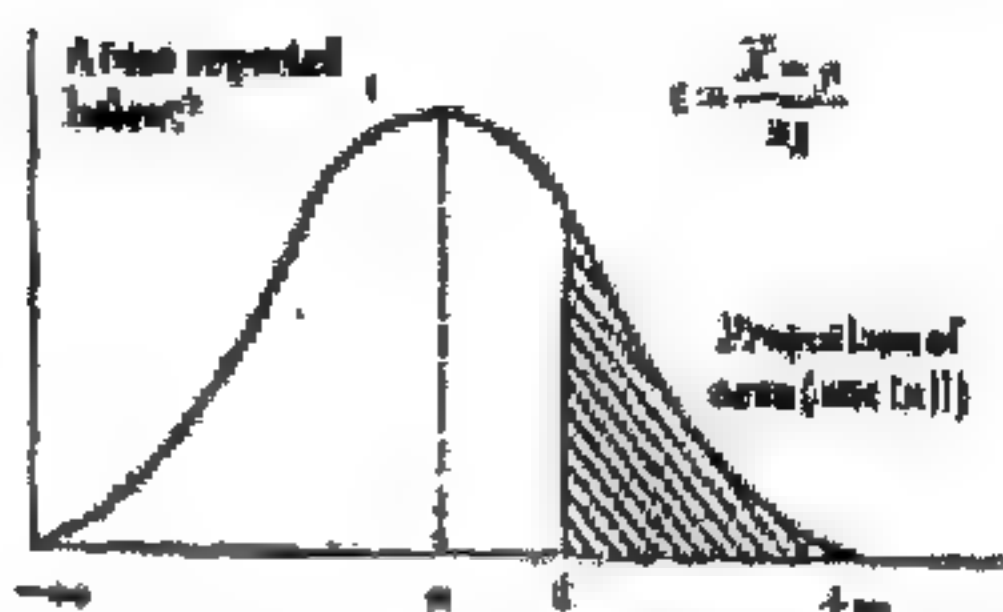
* Example: For $z = 1.96$, shaded area is 0.4750 out of the total area of 1.0000.

جدول توزيع T

المجدول أدناه يعطى قيمة t

المجاورة لمساحة المنطقة الواقعة تحتها α

Proportions of Area
for the t Distributions



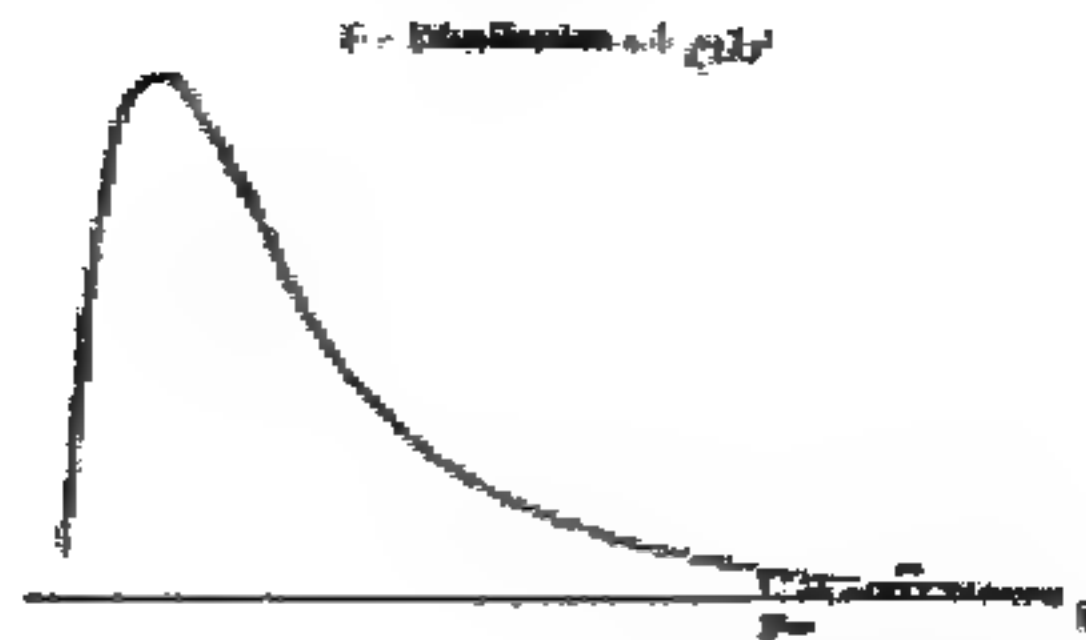
df	0.10	0.05	0.025	0.01	0.005
1	1.078	1.314	1.638	2.101	2.306
2	1.078	1.314	1.638	2.101	2.306
3	1.078	1.314	1.638	2.101	2.306
4	1.078	1.314	1.638	2.101	2.306
5	1.078	1.314	1.638	2.101	2.306
6	1.078	1.314	1.638	2.101	2.306
7	1.078	1.314	1.638	2.101	2.306
8	1.078	1.314	1.638	2.101	2.306
9	1.078	1.314	1.638	2.101	2.306
10	1.078	1.314	1.638	2.101	2.306
11	1.078	1.314	1.638	2.101	2.306
12	1.078	1.314	1.638	2.101	2.306
13	1.078	1.314	1.638	2.101	2.306
14	1.078	1.314	1.638	2.101	2.306
15	1.078	1.314	1.638	2.101	2.306
16	1.078	1.314	1.638	2.101	2.306
17	1.078	1.314	1.638	2.101	2.306

df	0.10	0.05	0.025	0.01	0.005
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.833
22	1.321	1.717	2.074	2.509	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.053	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
40	1.303	1.684	2.021	2.423	2.704
60	1.296	1.671	2.000	2.390	2.660
120	1.289	1.658	1.980	2.355	2.617
∞	1.282	1.645	1.960	2.326	2.576

* Example : For the shaded area to represent 0.05 of the area of t-Distribution with 10 degrees of freedom is 1.812.

Source: From Table III of Fisher and Yates, Statistical Tables for Biological, Agricultural and Medical Research, 6th, 1974, published by Longman Group Ltd., London (previously published by Oliver & Boyd, Edinburgh), by permission of the authors and publishers.

جدول توزيع F



للمعنى الأول $\alpha = 0.05$ والمعنى الثاني $\alpha = 0.01$

مثلي $F_{0.05}(1, 12) = 7.97$ و $F_{0.01}(1, 12) = 7.66$

درجات حرية المقام	درجات حرية البسط											
	1	2	3	4	5	6	7	8	9	10	11	12
1	161	199	216	225	230	234	237	239	241	242	243	244
2	4.052	4.999	5.403	5.625	5.764	5.850	5.928	5.981	6.022	6.056	6.082	6.106
3	18.51	19.00	19.16	19.25	19.30	19.33	19.36	19.37	19.38	19.39	19.40	19.41
4	98.49	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.41	99.42
5	10.13	9.58	9.28	9.12	9.01	8.94	8.89	8.84	8.81	8.78	8.76	8.74
6	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.34	27.23	27.13	27.05
7	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.93	5.91
8	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.54	14.45	14.37
9	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.78	4.74	4.70	4.68
10	16.26	13.27	12.06	11.39	10.97	10.67	10.45	10.28	10.15	10.05	9.96	9.89
11	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00
12	13.74	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.79	7.72
13	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.63	3.60	3.57
14	12.25	9.55	8.45	7.85	7.46	7.19	7.00	6.84	6.71	6.62	6.54	6.47
15	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.34	3.31	3.28
16	11.26	8.65	7.59	7.01	6.63	6.37	6.19	6.03	5.91	5.82	5.74	5.67
17	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.13	3.10	3.07
18	10.56	8.02	6.99	6.42	6.06	5.80	5.62	5.47	5.35	5.26	5.18	5.11
19	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.97	2.94	2.91
20	10.04	7.56	6.55	5.99	5.64	5.39	5.21	5.06	4.95	4.85	4.78	4.71
21	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.86	2.82	2.79
22	9.65	7.20	6.22	5.67	5.32	5.07	4.88	4.74	4.63	4.54	4.46	4.40
23	4.75	3.88	3.49	3.26	3.11	3.00	2.92	2.85	2.80	2.76	2.73	2.69
24	9.33	6.93	5.95	5.41	5.06	4.82	4.63	4.50	4.39	4.30	4.22	4.16
25	4.67	3.80	3.41	3.18	3.02	2.92	2.84	2.77	2.72	2.67	2.63	2.60
26	9.07	6.70	5.74	5.20	4.86	4.62	4.44	4.30	4.19	4.10	4.02	3.96
27	4.60	3.74	3.34	3.11	2.96	2.85	2.77	2.70	2.65	2.60	2.56	2.53
28	8.86	6.51	5.56	5.03	4.69	4.46	4.28	4.14	4.03	3.94	3.86	3.80
29	4.54	3.68	3.29	3.06	2.90	2.79	2.70	2.64	2.59	2.55	2.51	2.48
30	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.73	3.67
31	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.45	2.42
32	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.61	3.55
33	4.45	3.59	3.20	2.96	2.81	2.70	2.62	2.55	2.50	2.45	2.41	2.38
34	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.52	3.45
35	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34
36	8.23	6.01	5.09	4.58	4.25	4.01	3.85	3.71	3.60	3.51	3.44	3.37

جدول توزيع χ^2

Chi-Square Distribution

توزيع كاي تربيع

الجدول أعلاه يعطي قيمة χ^2 المقابلة للمساحة المظللة وقيمتها α

درجات الحرية	المساحة المظللة						
	.99	.98	.95	.90	.80	.70	.50
1	.03157	.03628	.00393	0.0158	.0642	.148	.455
2	.0201	.0404	.103	.211	.446	.713	1.386
3	.115	.185	.352	.584	1.005	1.424	2.366
4	.297	.429	.711	1.064	1.649	2.195	3.357
5	.554	.752	1.145	1.610	2.343	3.000	4.351
6	.872	1.134	1.635	2.204	3.070	3.828	5.348
7	1.239	1.564	2.167	2.833	3.822	4.671	6.346
8	1.646	2.032	2.733	3.490	4.594	5.527	7.344
9	2.088	2.532	3.325	4.168	5.390	6.393	8.343
10	2.558	3.059	3.940	4.865	6.179	7.267	9.342
11	3.053	3.609	4.575	5.578	6.989	8.148	10.341
12	3.571	4.178	5.226	6.304	7.807	9.034	11.340
13	4.107	4.765	5.892	7.042	8.634	9.926	12.340
14	4.660	5.368	6.571	7.790	9.467	10.821	13.339
15	5.229	5.985	7.261	8.547	10.307	11.721	14.339
16	5.812	6.614	7.962	9.312	11.152	12.624	15.338
17	6.408	7.255	8.672	10.085	12.002	13.531	16.338
18	7.015	7.906	9.390	10.865	12.857	14.440	17.338
19	7.633	8.567	10.117	11.651	13.716	15.352	18.338
20	8.260	9.237	10.851	12.443	14.578	16.266	19.337
21	8.897	9.915	11.591	13.240	15.445	17.182	20.337
22	9.542	10.600	12.338	14.041	16.314	18.101	21.337
23	10.196	11.293	13.091	14.848	17.187	19.021	22.337
24	10.856	11.992	13.848	15.659	18.062	19.943	23.337
25	11.524	12.697	14.611	16.473	18.940	20.867	24.337
26	12.198	13.409	15.379	17.292	19.820	21.792	25.336
27	12.879	14.125	16.151	18.114	20.703	22.719	26.336
28	13.565	14.847	16.928	18.939	21.588	23.647	27.336
29	14.256	15.574	17.708	19.768	22.475	24.577	28.336
30	14.951	16.306	18.491	20.599	23.364	25.508	29.336

المصادر Bibliography

المصادر العربية Arabic Bibliography

1. باهي، مصطفى حسين، عبد الفتاح احمد (2006). الاحصاء التطبيقي باستخدام الحزم الجاهزة SPSS, SAT . مصر، القاهرة: مكتبة الانجلو المصرية.
2. البنداوي، عبد الحميد (1997). الإحصاء للعلوم الإدارية والتطبيقية. الأردن، عمان: دار الشروق للنشر والتوزيع.
3. البياتي، محمود (2005). تحليل البيانات الإحصائية باستخدام البرنامج الإحصائي SPSS. الأردن، عمان: دار الحامد للنشر والتوزيع.
4. ابو حطب، فؤاد (1962). القدرات العقلية (ط5). مصر، القاهرة: مكتبة الانجلو المصرية.
5. الزعبي، بلال محمد والطلافة، عباس (2000). النظام الإحصائي SPSS فهم وتحليل البيانات الإحصائية. الأردن، عمان: دار وائل للطباعة والنشر.
6. السيد، فؤاد البهي (1987). علم النفس الإحصائي وقياس العقل البشري. مصر، القاهرة: دار الفكر العربي.
7. عبد الجبار، توفيق (1983). التحليل الإحصائي في البحوث التربوية والنفسية والاجتماعية-الطرق اللامعلمية. الكويت: مؤسسة الكويت للتقدم العلمي. إدارة التأليف والترجمة.
8. عبد الخالق، احمد محمد (1987). الابعاد الاساسية للشخصية. مصر، الاسكندرية: دار المعرفة الجامعية.
9. علام، صلاح الدين (1993). الأساليب الإحصائية الاستدلالية البارامترية واللابارامترية في تحليل بيانات البحوث النفسية والتربوية. مصر، القاهرة: دار الفكر العربي.

10. عودة، احمد سليمان والخليلي، تحليل (1988). الإحصاء للباحث في التربية والعلوم الإنسانية. الأردن، عمان: دار الفرقان للنشر والتوزيع.
11. فتح الله، سعيد حسين (1988). مبادئ علم الإحصاء والطرق الإحصائية. الأردن، المفرق: الأكاديمية.
12. فراج، محمد انور (2002). المكونات العملية للتفكير الناقد لدى طلاب كليات التربية في ضوء بعض المتغيرات، رسالة دكتوراه غير منشورة، كلية التربية، جامعة الاسكندرية، الاسكندرية، مصر.
13. انقاسم، محمد علي (1987). أساليب الإحصاء التطبيقي. الكويت: المعهد العربي للتخطيط بالكويت، دار الشباب للنشر والترجمة.
14. الكيلاني، عبد الله وعدس، عبد الرحمن (1986). القياس والتقويم في علم النفس والتربية. ترجمة: روبرت ثورندايك واليزابيث هيغن، الأردن، عمان: مركز الكتب الأردني.
15. منصور عوض ، عزام صبري ، علي قوقزة (1999). علم الإحصاء الوصفي المبرمج. الأردن، عمان: دار صفاء للنشر.
16. موراي، ر. شيرجل (1977). سلسلة ملخصات شوم، دار ماجدوهيل للنشر.
17. الميزل، عبد الله فلاح (2006). الإحصاء الاستدلالي وتطبيقاته في الحاسوب باستخدام الرزم الإحصائية SPSS، الأردن، عمان: دار وائل للطباعة والنشر والتوزيع.
18. الميزل، عبد الله والغرايبة، عايش (1995). الإحصاء التربوي تطبيقات باستخدام الرزم الإحصائية للعلوم الاجتماعية.
19. السنجار، نبيل جمعه (2004). مهارات الحاسوب، الأردن، اربد: عالم الكتب الحديث للنشر والتوزيع.
20. السنجار، نبيل جمعه (2007). الإحصاء في التربية والعلوم الإنسانية مع تطبيقات برمجية SPSS ، الأردن، عمان: دار الحامد للنشر والتوزيع.
21. السنجار، نبيل جمعه (2010). القياس والتقويم مع تطبيقات برمجية SPSS ، الأردن، عمان: دار الحامد للنشر والتوزيع.

22. النجار، فايز جمعه، النجار، نبيل جمعه والزعبي، ماجد راضي (2013). *أساليب البحث العلمي: منظور تطبيقي*، الأردن، عمان: دار الحامد للنشر والتوزيع.
23. الهانسي، مختار محمود (1991). *مقدمة في طرق التحليل الإحصائي*. قسم الإحصاء والرياضة والتأمين - كلية التجارة - جامعة الإسكندرية.
24. الهانسي، مختار محمود (1984). *مقدمة في الإحصاء التحليلي*. بيروت: دار النهضة العربية للطباعة والنشر.
25. هكسي، تشارلز (1984). *المفاهيم الأساسية في تصميم التجارب*. تعريب: خماس، قيس سبع، العراق، بغداد: الجامعة المستنصرية.

المصادر الأجنبية English Bibliography

1. Albert K. Kurtz, Samuel T. Mayo (1979). *Statistical Methods in Education and Psychology*. New York: Springer-Verlag, New York Inc.
2. Andy, Field (2005). *Discovering Statistics Using SPSS*. (2nd ed.). London: Sage Publications Ltd, ECIY 1SP.
3. Berenson, L. Mark & David M. Levine. (1992). *Basic Business Statistics Concepts and Applications*, (5th ed.).
4. Bobko, Philip (2001), *Correlation and regression*, (2nd ed.). Thousand Oaks, CA: Sage Publications. Introductory text which includes coverage of range restriction, trivariate correlation.
5. Chen, P. Y. and P. M. Popovich (2002). *Correlation: Parametric and nonparametric measures*. Thousand Oaks, CA: Sage Publications.
6. Cochran, W. G. (1997). *Sampling Techniques* (3rd ed.). New York.
7. Cohen, Jacob (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum Assoc; ISBN: 0805802835.
8. Crawley, M. J. (2005). *Statistics An Introduction Using R*. England: John Wiley & Sons Ltd, West Sussex.
9. Crocker, Lind, and Algina, Janes .(1986). *Introduction to Classical and Modern Test Theory*. N. Y: Holt, Rineder and Winson
10. David, R. Anderson, Dennis J. Sweeney , Thomas A. Williams (2002). *Essentials of Modern Business Statistics with Microsoft Excel* (2nd ed.). South-Western Educational Publishing.

11. David, R. Anderson, Dennis, J. Sweeney , Thomas A. Williams (2006). ***Essentials of Modern Business Statistics*** (3rd ed.). South-Western College Publishing.
12. David, R. Anderson, Dennis, J. Sweeney , Thomas, A. Williams (2005). ***Statistics for Business and Economics*** (9th ed.). South-Western College Publishing.
13. David Anderson, Dennis J. Sweeney & Thomas Williams, (1981). ***Introduction to Statistics***, West Publishing Co.
14. Daniel, Wayne W. (1990). ***Applied Nonparametric Statistics***, (2nd ed.). PWS-Kent Publishing Company-Boston.
15. Daniel, Wayne W. (1995). ***Biostatistics A Foundation for Analysis in the Health Sciences*** (6th ed.). London: John Wiley and Sons Inc.
16. Deming, W. (1982). ***Applied Regression Analysis***. London: John Wiley and Sons Inc.
17. Dominick Salvator (1982). ***Theory and Problems of Statistics and Econometrics***, McGraw-Hill Book Co.
18. Draper N. & Smith H. (1990). ***Applied Regression Analysis***. London: John Wiley and Sons Inc.
19. Guilford, J.P (1961). ***Factorial analysis to Psychology***.
20. Glass, G. & Hopkins, K. D.(1984). ***Statistical Methods in Educational and psychology***(2nd ed.). Englewood Cliffs, New Jersey: Prentice-Hall.
21. Gary W. Heiman. (1992). ***Basic Statistics for the Behavioral Sciences***. Houghton Mifflin Company, Boston, U.S.A
22. Graber, S. B., Kristin E. Voelki, T. W. & Others .(1997). ***SPSS Guide to the New Statistical Analysis of Data***, New York: Springer.
23. Grant, E. & Leavenworth R., (1980). ***Statistical Quality Control***, (5th ed.). New York: McGraw-Hill.
24. Gravetter, F. O.(1988). ***Statistics for the behavioral sciences***. New York: West Publishing company.
25. Green, Samuel B. & Niel J. Salkind (1997). ***Using SPSS for Windows: Analyzing and Understanding Data***, Upper Saddle River. NJ: Prentice Hall.
26. Gregory, B., (1991). ***Introduction to Quality Management Assurance and Control***, (Macmillan International ed). New York: Maxwell.
27. Hays, W. L. (1980). ***Statistics for Social Sciences*** (3rd ed.). New York: Holt Rinehart and Winston.

28. Hays, W. L. (1989). **Statistics** (4th ed.). New York: Holt Rinehart and Winston.
29. Hogg, N. R. & Carag, T. A. (1995). **Introduction to Mathematical Statistics** (5th ed.). Prentice-Hall Inc.
30. Howitt, Dennis & Duncan, Cramer (1996). **A Guide to Computing Statistical with SPSS for Windows**, New York: Prentice-Hall/ Harvester Wheatsheaf.
31. Jaynes, E. T. & G. Larry Bretthorst (2003). **Probability Theory: The Logic of Science**. Cambridge University.
32. Johnson, Richard A. & Wichern, D. (1992). **Applied Multivariate Statistical** (3rd). New jersey: prencac-hall.
33. Juran, J. & Gryna, F. (1970). **Quality Planning and Analysis**. New Yourk: McGraw-Hill.
34. Kendall, Maurice and Jean Dickinson Gibbons (1990). **Rank Correlation Methods**. (5th ed.). NY: Oxford Univ Press; ISBN: 0195208374.
35. Kenneth, D. Hopkins, Julian C. Stanley & B. R. Hopkins .(1995). **Educational and Psychological Measurement and Evaluation**. (7th ed.). Allyn & Bacon, Needham Hights, USA.
36. Larson, Harold. J. (1982). **Introduction to Probability Theory and Statistical Inference** (3rd ed.). John Wiely & Sons Inc.
37. Larson, Harold. J. (1995). **Introduction to Probability**, Addison Wesley Publishing Company Inc.
38. Larry, J. Stephens (2005). **Schaum's Outline of Beginning Statistics** (2nd ed.). McGraw-Hill.
39. Mark, S. Aldenderfer & Roger, K. Blashfield (1991). **Cluster Analysis** (8th ed.). sage Publications - The International Professional Publishers, Eighth Printing.
40. (Manual) (1997). **SPSS Base 7.5 Application Guide**, SPSS Inc.
41. (Manual) (1997). **SPSS Base 7.5 for Windows User Guide**, SPSS Inc.
42. (Manual) (1994). **SPSS Advanced Statistics 6.1**, SPSS Inc.
43. Marija J. Norusis (1993). **SPSS for Windows Base System User's Guide Release 6.0 (Manual)**, SPSS Inc.
44. McClave, J. T. & Benson, G. P. (1991). **Statistics For Business and Economics** (5th ed.). San Fransisco Dellan.
45. Mendenhall W. & Sincich (1991). **Statistics For Engineering and the Sciences** (3rd ed.). New York: Macmillan Publishing Co.

46. Minium, E. W.(1978). ***Statistical Reasoning in psychology and Education*** (2nd ed.). New York: John Wiley and Sons.
47. Morris, H, (1983). ***Statistical Analysis for Decision Making***. (3rd ed.). USA: Harcourt Brace Jovanovich Inc.
48. Nelile, A. & Kennedy, J. (1974). ***Basic Statistical Methods for Engineers & Scientists***, London: Inter Text Books.
49. Parzen, E. (1960). ***Modern Probability theory and its application***, New York: John Wiley and Sons.
50. Siegel, S. (1956). ***Nonparametric statistics for the behavioral sciences***. NY: McGraw-Hill.





الإحصاء التحليلي مع تطبيقات برمجية SPSS



للنسخة
5330856



دار الحداثة للنشر والتوزيع

الأردن - عمان

هاتف: 5231081 فاكس: +96265235594

ص.ب: 366 عمان 11941 الأردن

E-mail: dar_alhamed@hotmail.com

E-mail: Daralhamed@yahoo.com

کتاب ترابوہ



Educational books

Telegram : t.me/edubook